L Number	Hits	Search Text	DB	Time stamp
1	0	liu-qingyun.in	USPAT;	2003/02/06 08:28
			US-PGPUB;	
			EPO; JPO;	
_ [DERWENT]
2	7	liu-qingyun.in.	USPAT;	2003/02/06 08:28
			US-PGPUB;	
			EPO; JPO;	
	_		DERWENT	
3	3	mcdonald-terrence-p.in.	USPAT;	2003/02/06 08:28
			US-PGPUB;	
i			EPO; JPO;	
	_		DERWENT	
4	2	hg51	USPAT;	2003/02/06 08:29
			US-PGPUB;	
			EPO; JPO;	
			DERWENT	

Welcome to STN International! Enter x:x

LOGINID:ssspta1649jxm

PASSWORD:

TERMINAL (ENTER 1, 2, 3, OR ?):2

```
Welcome to STN International
 NEWS
                 Web Page URLs for STN Seminar Schedule - N. America
 NEWS
      2 Apr 08
                 "Ask CAS" for self-help around the clock
 NEWS
         Jun 03
                 New e-mail delivery for search results now available
      4 Aug 08
 NEWS
                 PHARMAMarketLetter(PHARMAML) - new on STN
      5 Aug 19 Aquatic Toxicity Information Retrieval (AQUIRE)
 NEWS
                 now available on STN
 NEWS
      6 Aug 26 Sequence searching in REGISTRY enhanced
     7 Sep 03 JAPIO has been reloaded and enhanced
 NEWS
 NEWS 8 Sep 16 Experimental properties added to the REGISTRY file
NEWS 9 Sep 16 CA Section Thesaurus available in CAPLUS and CA
NEWS 10 Oct 01 CASREACT Enriched with Reactions from 1907 to 1985
NEWS 11 Oct 24 BEILSTEIN adds new search fields
NEWS 12 Oct 24 Nutraceuticals International (NUTRACEUT) now available on STN
NEWS 13 Nov 18 DKILIT has been renamed APOLLIT
NEWS 14 Nov 25 More calculated properties added to REGISTRY
NEWS 15 Dec 04 CSA files on STN
NEWS 16 Dec 17 PCTFULL now covers WP/PCT Applications from 1978 to date
NEWS 17 Dec 17
                 TOXCENTER enhanced with additional content
NEWS 18 Dec 17
                 Adis Clinical Trials Insight now available on STN
NEWS 19 Jan 29 Simultaneous left and right truncation added to COMPENDEX,
                 ENERGY, INSPEC
NEWS 20 Feb 13 CANCERLIT is no longer being updated
NEWS 21 Feb 24 METADEX enhancements
NEWS 22 Feb 24 PCTGEN now available on STN
NEWS 23 Feb 24
                TEMA now available on STN
NEWS 24 Feb 26 NTIS now allows simultaneous left and right truncation
NEWS 25 Feb 26 PCTFULL now contains images
NEWS 26 Mar 04 SDI PACKAGE for monthly delivery of multifile SDI results
NEWS 27 Mar 20 EVENTLINE will be removed from STN
NEWS 28 Mar 24
                 PATDPAFULL now available on STN
NEWS 29
        Mar 24
                Additional information for trade-named substances without
                 structures available in REGISTRY
NEWS 30
        Apr 11
                Display formats in DGENE enhanced
NEWS 31
        Apr 14
                MEDLINE Reload
NEWS 32
        Apr 17
                 Polymer searching in REGISTRY enhanced
NEWS 33
                Indexing from 1947 to 1956 being added to records in CA/CAPLUS
        Apr 21
NEWS 34
        Apr 21
                New current-awareness alert (SDI) frequency in
                WPIDS/WPINDEX/WPIX
NEWS 35
                RDISCLOSURE now available on STN
        Apr 28
NEWS 36 May 05
                Pharmacokinetic information and systematic chemical names
                added to PHAR
NEWS 37
        May 15
                MEDLINE file segment of TOXCENTER reloaded
NEWS 38
        May 15
                Supporter information for ENCOMPPAT and ENCOMPLIT updated
NEWS 39
        May 16
                CHEMREACT will be removed from STN
NEWS 40
        May 19
                Simultaneous left and right truncation added to WSCA
        May 19 RAPRA enhanced with new search field, simultaneous left and
NEWS 41
                right truncation
```

NEWS EXPRESS April 4 CURRENT WINDOWS VERSION IS V6.01a, CURRENT MACINTOSH VERSION IS V6.0b(ENG) AND V6.0Jb(JP), AND CURRENT DISCOVER FILE IS DATED 01 APRIL 2003 NEWS HOURS STN Operating Hours Plus Help Desk Availability NEWS INTER General Internet Information NEWS LOGIN Welcome Banner and News Items NEWS PHONE Direct Dial and Telecommunication Network Access to STN NEWS WWW CAS World Wide Web Site (general information) Enter NEWS followed by the item number or name to see news on that specific topic. All use of STN is subject to the provisions of the STN Customer agreement. Please note that this agreement limits use to scientific research. Use for software development or design or implementation of commercial gateways or other similar uses is prohibited and may result in loss of user privileges and other penalties. * * * * * * * * * * * * * * STN Columbus FILE 'HOME' ENTERED AT 09:57:22 ON 28 MAY 2003 => file medline biosis embase caplus COST IN U.S. DOLLARS SINCE FILE TOTAL ENTRY SESSION FULL ESTIMATED COST 0.21 0.21 FILE 'MEDLINE' ENTERED AT 09:57:38 ON 28 MAY 2003 FILE 'BIOSIS' ENTERED AT 09:57:38 ON 28 MAY 2003 COPYRIGHT (C) 2003 BIOLOGICAL ABSTRACTS INC. (R) FILE 'EMBASE' ENTERED AT 09:57:38 ON 28 MAY 2003 COPYRIGHT (C) 2003 Elsevier Science B.V. All rights reserved. FILE 'CAPLUS' ENTERED AT 09:57:38 ON 28 MAY 2003 USE IS SUBJECT TO THE TERMS OF YOUR STN CUSTOMER AGREEMENT. PLEASE SEE "HELP USAGETERMS" FOR DETAILS. COPYRIGHT (C) 2003 AMERICAN CHEMICAL SOCIETY (ACS) => s liu qingyun /au 91 LIU QINGYUN => s mcdonald terrence /au 7 MCDONALD TERRENCE => s l1 and l2 2 L1 AND L2 => d l3 total ibib ANSWER 1 OF 2 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC. ACCESSION NUMBER: 1999:204159 BIOSIS DOCUMENT NUMBER: PREV199900204159 TITLE: Cloning of a novel G-protein-coupled receptor GPR 51 resembling GABAB receptors expressed predominantly in nervous tissues and mapped proximal to the hereditary sensory neuropathy type 1 locus on chromosome 9. AUTHOR (S): Ng, Gordon Y. K. (1); McDonald, Terrence; Bonnert, Tim; Rigby, Michael; Heavens, Robert; Whiting, Paul; Chateauneuf, Anne; Coulombe, Nathalie; Kargman, Stacia; Caskey, Thomas; Evans, Jilly; O'Neill, Gary P.; Liu, Qingyun (1) Department of Biochemistry and Molecular Biology, Merck CORPORATE SOURCE:

Frosst Center for Therapeutic Research, 16711 TransCanada

Highway, Kirkland, PQ, H9H 3L1 Canada

Genomics, (March 15, 1999) Vol. 56, No. 3, pp. 288-295. SOURCE:

ISSN: 0888-7543.

DOCUMENT TYPE:

Article English

LANGUAGE: SUMMARY LANGUAGE:

English

ANSWER 2 OF 2 CAPLUS COPYRIGHT 2003 ACS ACCESSION NUMBER:

1999:191160 CAPLUS

DOCUMENT NUMBER:

131:98163

TITLE:

Cloning of a Novel G-Protein-Coupled Receptor GPR 51 Resembling GABAB Receptors Expressed Predominantly in Nervous Tissues and Mapped Proximal to the Hereditary

Sensory Neuropathy Type 1 Locus on Chromosome 9

AUTHOR (S):

Ng, Gordon Y. K.; McDonald, Terrence;

Bonnert, Tim; Rigby, Michael; Heavens, Robert;

Whiting, Paul; Chateauneuf, Anne; Coulombe, Nathalie;

Kargman, Stacia; Caskey, Thomas; Evans, Jilly;

O'Neill, Gary P.; Liu, Qingyun

CORPORATE SOURCE:

Department of Biochemistry and Molecular Biology,

Merck Frosst Center for Therapeutic Research,

Kirkland, QC, H9H 3L1, Can. Genomics (1999), 56(3), 288-295

CODEN: GNMCEP; ISSN: 0888-7543

PUBLISHER:

SOURCE:

Academic Press

DOCUMENT TYPE:

Journal

LANGUAGE:

English 24

REFERENCE COUNT:

THERE ARE 24 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

=> s hg51 (s) protein (s) coupled (s) receptor 3 HG51 (S) PROTEIN (S) COUPLED (S) RECEPTOR

=> d l4 total ibib kwic

ANSWER 1 OF 3 CAPLUS COPYRIGHT 2003 ACS ACCESSION NUMBER:

DOCUMENT NUMBER:

2001:232346 CAPLUS

135:314242

TITLE:

Characterization of a novel human opsin gene with wide tissue expression and identification of embedded and

flanking genes on chromosome 1q43

AUTHOR (S):

Halford, Stephanie; Freedman, Melanie S.; Bellingham, James; Inglis, Suzanne L.; Poopalasundaram, Subathra; Soni, Bobby G.; Foster, Russell G.; Hunt, David M.

CORPORATE SOURCE:

Department of Molecular Genetics, Institute of Ophthalmology, University College London, London, EC1V

9EL, UK

SOURCE:

Genomics (2001), 72(2), 203-208 CODEN: GNMCEP; ISSN: 0888-7543

PUBLISHER:

Academic Press

DOCUMENT TYPE: LANGUAGE:

Journal

English

REFERENCE COUNT:

THERE ARE 25 CITED REFERENCES AVAILABLE FOR THIS 25 RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

ΙT 273191-92-7, G protein-coupled receptor HG51 (human)

RL: BSU (Biological study, unclassified); PRP (Properties); BIOL (Biological study)

(amino acid sequence; characterization of a novel human opsin gene with wide tissue expression and identification of embedded and flanking genes on chromosome 1q43)

ANSWER 2 OF 3 CAPLUS COPYRIGHT 2003 ACS L4ACCESSION NUMBER: 2000:513802 CAPLUS DOCUMENT NUMBER: 133:130801 TITLE: Cloning of a novel human G-protein-coupled receptor (GPCR) -17723 receptor cDNA and its therapeutic use INVENTOR(S): Glucksmann, Maria Alexandra PATENT ASSIGNEE(S): Millennium Pharmaceuticals, Inc., USA SOURCE: PCT Int. Appl., 79 pp. CODEN: PIXXD2 DOCUMENT TYPE: Patent LANGUAGE: English FAMILY ACC. NUM. COUNT: PATENT INFORMATION: PATENT NO. KIND DATE APPLICATION NO. DATE
WO 2000043513 A1 20000727 WO 2000-US1592 20000121 W: AE, AL, AM, AT, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, CZ, DE, DE, DK, DK, DM, EE, EE, ES, FI, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM RW: GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG PRIORITY APPLN. INFO.: US 1999-234923 A 19990121 REFERENCE COUNT: THERE ARE 6 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT 273191-92-7P, G Protein-coupled receptor HG51 (human) RL: BPN (Biosynthetic preparation); BPR (Biological process); BSU (Biological study, unclassified); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); PREP (Preparation); PROC (Process); USES (Uses) (amino acid sequence; cloning of novel human G-proteincoupled receptor (GPCR) -17723 receptor cDNA and its therapeutic use) ANSWER 3 OF 3 CAPLUS COPYRIGHT 2003 ACS ACCESSION NUMBER: 2000:368396 CAPLUS DOCUMENT NUMBER: 133:27864 TITLE: Human G protein-coupled receptor HG51, its sequence, cDNA encoding it, recombinant production and use in methods designed to identify agonists and/or antagonists INVENTOR(S): Liu, Qingyun; McDonald, Terrence P. PATENT ASSIGNEE(S): Merck & Co., Inc., USA SOURCE: PCT Int. Appl., 68 pp. CODEN: PIXXD2 DOCUMENT TYPE: Patent LANGUAGE: English FAMILY ACC. NUM. COUNT: 1 PATENT INFORMATION: PATENT NO. KIND DATE APPLICATION NO. DATE ---------WO 2000031108 A1 20000602 WO 1999-US27305 19991118 W: CA, JP, US RW: AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE EP 1133515 A1 20010919 EP 1999-962792 19991118 R: AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, PRIORITY APPLN. INFO.: US 1998-109717P P 19981124

WO 1999-US27305 W 19991118

REFERENCE COUNT: 4 THERE ARE 4 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

TI Human G protein-coupled receptor

HG51, its sequence, cDNA encoding it, recombinant production and use in methods designed to identify agonists and/or antagonists

The invention provides a cDNA mol. encoding a human G-protein-AΒ coupled receptor, HG51, as well as the receptor encoded by the cDNA mol. The invention also provides an expression vector (eukaryotic or prokaryotic) contg. HG51 cDNA mols. and host cells transformed with said vector, used for the recombinant prodn. of HG51. The invention further provides anti-HG51 antibodies. Still further, the invention provides methods for identifying substances that bind and/or modulate HG51, which include potential agonists and/or antagonists of HG51. The methods are cell based whereby an expression vector contg. polynucleotides encoding HG51 is transfected into a host cell, allowing for the recombinant prodn. of HG51 prior to addn. of the test substance. Finally, the invention provides the cDNA sequence, as well as the corresponding amino acid sequence of human HG51. The human HG51 receptor was shown to have sequence homol. to the rhodopsin subfamily of G protein-coupled receptors.

ST cDNA sequence human G protein coupled receptor HG51; recombinant prodn human G protein coupled receptor HG51; agonist human G protein coupled receptor HG51 method identification; antagonist human G protein coupled receptor HG51 method identification

IT G protein-coupled receptors

RL: BPN (Biosynthetic preparation); BPR (Biological process); BSU (Biological study, unclassified); BUU (Biological use, unclassified); PRP (Properties); BIOL (Biological study); PREP (Preparation); PROC (Process); USES (Uses)

(HG51; human G protein-coupled receptor HG51, its sequence, cDNA encoding it, recombinant prodn. and use in methods designed to identify agonists and/or antagonists)

IT Antibodies

RL: BSU (Biological study, unclassified); BIOL (Biological study) (antibodies specific for human G-protein-coupled receptor HG51)

IT Cell membrane

(contains recombinant HG51; human G protein-coupled receptor HG51, its sequence, cDNA encoding it, recombinant prodn. and use in methods designed to identify agonists and/or antagonists)

IT Genetic vectors

(eukaryotic or prokaryotic; human G protein-coupled receptor HG51, its sequence, cDNA encoding it, recombinant prodn. and use in methods designed to identify agonists and/or antagonists)

IT Molecular cloning

Transformation, genetic

(human G protein-coupled receptor

HG51, its sequence, cDNA encoding it, recombinant prodn. and
use in methods designed to identify agonists and/or antagonists)
Ligands

IT Ligands
RL: BPR (Biological process); BSU (Biological study, unclassified); BUU
(Biological use, unclassified); BIOL (Biological study); PROC (Process);
USES (Uses)

(of HG51; human G protein-coupled receptor HG51, its sequence, cDNA encoding it, recombinant prodn. and use in methods designed to identify agonists and/or antagonists)

cDNA sequences IT(of cDNA encoding human G protein-coupled receptor HG51) ITProtein sequences (of human G protein-coupled receptor 273191-92-7P, G Protein-coupled receptor ITHG51 (human) RL: BPN (Biosynthetic preparation); BPR (Biological process); BSU (Biological study, unclassified); BUU (Biological use, unclassified); PRP (Properties); BIOL (Biological study); PREP (Preparation); PROC (Process); USES (Uses) (amino acid sequence; human G protein-coupled receptor HG51, its sequence, cDNA encoding it, recombinant prodn. and use in methods designed to identify agonists and/or antagonists) IT273192-12-4D, subfragments are claimed RL: BUU (Biological use, unclassified); PRP (Properties); BIOL (Biological study); USES (Uses) (nucleotide sequence; cDNA mol. encoding human G-proteincoupled receptor HG51, its sequence and use in recombinant prodn. of HG51) 202544-98-7, 15: $\overline{\text{PN}}$: WO0031108 SEQID: 3 unclaimed DNA IT214908-89-1 214908-90-4 273192-88-4, 3: PN: WO0031108 SEQID: 4 unclaimed DNA 273192-89-5, 4: PN: WO0031108 SEQID: 5 unclaimed DNA 273192-90-8, 5: PN: WO0031108 SEQID: 6 unclaimed DNA 273192-91-9, 6: PN: WO0031108 SEQID: 7 unclaimed DNA 273192-92-0, 7: PN: WO0031108 SEQID: 8 unclaimed DNA 273192-93-1, 8: PN: WO0031108 SEQID: 9 unclaimed DNA 273192-94-2, 9: PN: WO0031108 SEQID: 10 unclaimed DNA 290797-99-8 RL: PRP (Properties) (unclaimed nucleotide sequence; human G proteincoupled receptor HG51, its sequence, cDNA encoding it, recombinant prodn. and use in methods designed to identify agonists and/or antagonists) ΙT 93050-34-1, Rhodopsin (human protein moiety reduced) RL: PRP (Properties) (unclaimed protein sequence; human G proteincoupled receptor HG51, its sequence, cDNA encoding it, recombinant prodn. and use in methods designed to identify agonists and/or antagonists)

SEQ ID NO: 1

```
Result
                 Query
    No. Score Match Length DB ID
                                                            Description
  -----
                                                           -----
     1 1470.4 95.7 2533 9 AF303588
                                                            AF303588 Homo sapi
                       2110 9 AF140242
      2 1380.8
                89.8
                                                            AF140242 Homo sapi
     3
          943.6
                 61.4
                       1737 10 AF140241
                                                            AF140241 Mus muscu
                 33.5 71872 9 AL133390
27.6 5024 6 AX281720
  C
     4
          515.4
                                                           AL133390 Human DNA
  C
     5
          424
                 27.6
                                                           AX281720 Sequence
  C
     6
          422.4
                27.5
                        5000 6 A68713
                                                            A68713 Sequence 5
  C
     7
          422.4
                 27.5
                        5000 6 AR106624
                                                           AR106624 Sequence
  С
     8
          422.4
                 27.5
                        5000 9 AF056032
449 6 AX013137
                                                           AF056032 Homo sapi
          400.4
                 26.1
                                                           AX013137 Sequence
     10
          358
                 23.3
                         512 6 AX062411
                                                           AX062411 Sequence
  С
    11
         317.2
                 20.6
                       8303 6 AX345325
                                                           AX345325 Sequence
    12
         299.4
                 19.5
                        8303 6 AX345324
                                                           AX345324 Sequence
RESULT
AF303588
LOCUS
            AF303588
                                     2533 bp
                                               mRNA
                                                       linear
                                                                PRI 17-APR-2001
DEFINITION Homo sapiens panopsin (OPN3) mRNA, complete cds.
            AF303588
VERSION
            AF303588.1 GI:13649589
KEYWORDS
SOURCE
            human.
  ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
            1 (bases 1 to 2533)
  AUTHORS
            Halford, S., Freedman, M.S., Bellingham, J., Inglis, S.L.,
            Poopalasundaram, S., Soni, B.G., Foster, R.G. and Hunt, D.M. Characterization of a novel human opsin gene with wide tissue
  TITLE
            expression and identification of embedded and flanking genes on
            chromosome 1q43
  JOURNAL
            Genomics 72 (2), 203-208 (2001)
  MEDLINE
            21295039
REFERENCE
            2 (bases 1 to 2533)
           Halford, S., Bellingham, J., Freedman, M.S., Inglis, S.L.,
  AUTHORS
            Poopalasundaram, S., Foster, R. and Hunt, D.M.
  TITLE
            Direct Submission
  JOURNAL
            Submitted (05-SEP-2000) Molecular Genetics, Institute of
            Ophthalmology, 11-43 Bath Street, London EC1V 9EL, UK
FEATURES
                    Location/Qualifiers
     source
                    1. .2533
                    /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /chromosome="1"
                     /map="1q43"
    gene
                    1. .2533
                     /gene="OPN3"
    CDS
                    103. .1311
                     /gene="OPN3"
                    /note="multiple tissue opsin"
                    /codon_start=1
                    /product="panopsin"
                    /protein_id="AAK37447.1"
                    /db xref="GI:13649590"
                    translation="MYSGNRSGGHGYWDGGGAAGAEGPAPAGTLSPAPLFSPGTYERL/
                    {\tt ALLLGSIGLLGVGNNLLVLYYKFQRLRTPTHLLLVNISLSDLLVSLFGVTFTFVSC}
                    LRNGWVWDTVGCVWDGFSGSLFGIVSIATLTVLAYERYIRVVHARVINFSWAWRAITY
                    IWLYSLAWAGAPLLGWNRYILDVHGLGCTVDWKSKDANDSSFVLFLFLGCLVVPLGVI
                    AHCYGHILYSIRMLRCVEDLQTIQVIKILKYEKKLAKMCFLMIFTFLVCWMPYIVICF
                    LVVNGHGHLVTPTISIVSYLFAKSNTVYNPVIYVFMIRKFRRSLLQLLCLRLLRCQRP
                    AKDLPAAGSEMQIRPIVMSQKDGDRPKKKVTFNSSSIIFIITSDESLSVDDSDKTNGS
                    KVDVIQVRPL"
```

Query Match 95.7%; Score 1470.4; DB 9; Length 2533; Best Local Similarity 99.6%; Pred. No. 7.8e-229; Matches 1474; Conservative 0; Mismatches Indels 0; Gaps Qу 1 CGGACGCGTGGGCGGAGCCCCAGCCCACCCAGTGCGGAGCGCGCGGAGCCCCGCCGC 60 Db Qy Db 61 AAGCTGAGCGCCTCCGCCCGCCAGGCGCCCGGCCCGGGCCATGTACTCGGGGAACCGC 120 Qy Db 238 gcggggacactgagccccgcgcccctcttcagccccggcacctacgagcgcctggcgctg 297 Qу ក្រុម្រីក្រុមប្រជាជ្រាកស៊ូរប៉ុស្តែអាវិកម្មាធិប្រកួរប្រកួតអាចក្រុមប្រជាជាក្រុមប្រជាជាការប្រ 181 GCGGGGACACTGAGCCCCGCGCCCCTCTTCAGCCCCGGCACCTACGAGCGCCTGGCGCTG 240 Db Qу 298 ctgctgggctccattgggctgctgggcgtcggcaacaacctgctggtgctcgtcctctac 357 241 CTGCTGGGCTCCATTGGGCTGCTGGGCGTCGGCAACAACCTGCTGGTGCTCGTCCTCTAC 300 Db 358 tacaagttccagcggctccgcactccactcacctcctcctggtcaacatcagcctcagc 417 Qy 301 TACAAGTTCCAGCGGCTCCGCACTCCCACTCCTCCTCGTCAACATCAGCCTCAGC 360 Db Ov Db 478 tgggtgtgggacaccgtgggctgcgtgtgggacgggtttagcggcagcctcttcgggatt 537 Qу Db 421 TGGGTGTGGGACACCGTGGGCTGCGTGTGGGACGGGTTTAGCGGCAGCCTCTTCGGGATT 480 538 gtttccattgccaccctaaccgtgctggcctatgaacgttacattcgcgtggtccatgcc 597 Qу 481 GTTTCCATTGCCACCCTAACCGTGCTGGCCTATGAACGTTACATTCGCGTGGTCCATGCC 540 Qу agagtgatcaatttttcctgggcctggagggccattacctacatctggctctactcactg 657 541 AGAGTGATCAATTTTTCCTGGGCCTGGAGGGCCATTACCTACATCTGGCTCTACTCACTG 600 Db Qу 658 gcgtgggcaggagcacctctcctgggatggaacaggtacatcctggacgtacacggacta 717 Db 601 GCGTGGGCAGGAGCACCTCTCCTGGGATGGAACAGGTACATCCTGGACGTACACGGACTA 660 718 ggctgcactgtggactggaaatccaaggatgccaacgattcctcctttgtgcttttctta 777 Qу Db 661 GGCTGCACTGTGGACTGGAAATCCAAGGATGCCAACGATTCCTCCTTTGTGCTTTTCTTA 720 778 tttcttggctgcctggtggtgcccctgggtgtcatagcccattgctatggccatattcta 837 Qу TTTCTTGGCTGCCTGGTGCCCCTGGGTGTCATAGCCCATTGCTATGGCCATATTCTA 780 Db 838 tattccattcgaatgcttcgttgtgtggaagatcttcagacaattcaagtgatcaagatt 897 Qу Db 781 TATTCCATTCGAATGCTTCGTTGTGTGGAAGATCTTCAGACAATTCAAGTGATCAAGATT 840 Qу 898 ttaaaatatgaaaagaaactggccaaaatgtgctttttaatgatattcaccttcctggtc 957 Db 841 TTAAAATATGAAAAGAAACTGGCCAAAATGTGCTTTTTAATGATATTCACCTTCCTGGTC 900 Oy 958 tgttggatgccttatatcgtgatctgcttcttggtggttaatggtcatggtcacctggtc 1017 901 TGTTGGATGCCTTATATCGTGATCTGCTTCTTGGTGGTTAATGGTCATGGTCACCTGGTC 960

0;

```
1018 actccaacaatatctattgtttcgtacctctttgctaaatcgaacactgtatacaatcca 1077
         Db
      961 ACTCCAACAATATCTATTGTTTCGTACCTCTTTGCTAAATCGAACACTGTATACAATCCA 1020
     1078 gtgatttatgtetteatgateagaaagtttegaagateeettttgeagettetgtgeete 1137
 Qу
         Db
     1021 GTGATTTATGTCTTCATGATCAGAAAGTTTCGAAGATCCCTTTTGCAGCTTCTGTGCCTC 1080
    1138 cgactgctgaggtgccagaggcctgctaaagacctaccagcagctggaagtgaaatgcag 1197
 Qу
         Db
    1081 CGACTGCTGAGGTGCCAGAGGCCTGCTAAAGACCTACCAGCAGCTGGAAGTGAAATGCAG 1140
    1198 atcagacccattgtgatgtcacagaaagatggggacaggccaaagaaaaagtgactttc 1257
 Qу
         Db
    1141 ATCAGACCCATTGTGATGTCACAGAAAGATGGGGGACAGGCCAAAGAAAAAGTGACTTTC 1200
    1258 aactettettecateatttttateateaceagtgatgaateaetgteagttgaegaeage 1317
Ov
         1201 AACTCTTCCTTCCATCATTTTTATCATCACCAGTGATGAATCACTGTCAGTTGACGACAGC 1260
Ov
    1318 gacaaaaccaatgggtccaaagttgatgtaatccaagttcgtcctttgtaggaatgaaga 1377
         1261 GACAAAACCAATGGGTCCAAAGTTGATGTAATCCAAGTTCGTCCTTTGTAGGAATGAAGA 1320
Qу
    1378 atggcaacgaaagatggggccttaaattggatgccacttttggactttcatcataagaag 1437
         1321 ATGGCAACGAAAGATGGGGCCTTAAATTGGATGCCACTTTTGGACTTTCATCATAAGAAG 1380
Qу
    1438 tgtctggaatacccgttctatgtaatatcaacagaaccttgtggtccagcaggaaatccg 1497
         Ďb
    1381 TGTCTGGAATACCCGTTCTATGTAATATCAACAGAACCTTGTGGTCCAGCAGGAAATCCG 1440
Qу
    1498 aattgcccatatgctcttgggcctcaggaagaggttgaac 1537
        Db
    1441 AATTGCCCATATGCTCTTGGGCCTCAGGAAGAGGTTGAAC 1480
RESULT 2
AF140242
                             2110 bp
                                      mRNA
                                             linear PRI 27-MAY-1999
DEFINITION Homo sapiens encephalopsin mRNA, complete cds.
ACCESSION
         AF140242
         AF140242.1 GI:4894951
VERSION
KEYWORDS
SOURCE
         human.
 ORGANISM
         Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
         Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
         1 (bases 1 to 2110)
 AUTHORS
         Blackshaw, S. and Snyder, S.H.
 TITLE
         Encephalopsin: A novel mammalian extraretinal opsin discretely
         localized in the brain
 JOURNAL.
         J. Neurosci. 19 (10), 3681-3690 (1999)
 MEDLINE
         99252448
  PUBMED
         10234000
REFERENCE
         2 (bases 1 to 2110)
 AUTHORS
         Blackshaw, S. and Snyder, S.H.
 TITLE
         Direct Submission
 JOURNAL
         Submitted (02-APR-1999) Genetics, Harvard Medical School, 200
         Longwood Ave., Boston, MA 02115, USA
FEATURES
                Location/Qualifiers
   source
                 1. .2110
                 /organism="Homo sapiens"
                 /db xref="taxon:9606"
   CDS
                 56. .1267
                 /note="extraretinal opsin"
                 /codon_start=1
                 /product="encephalopsin"
                 /protein_id="AAD32671.1"
                 /db_xref="GI:4894952"
```

Qу

/translation="mysgnrsgghgywdgggaagaegpapagtlspaplfspgtyerl
ALLLGSIGLLGVGNNLLVLVLYYKFQRLRTPTHLLLVNISLSDLLVSLFGVTFTFVSC
LRNGWVWDTVGCVWDGFSGSLFGIVSIATLTVLAYERYIRVVHARVINFSWAWRAITY
IWLYSLAWAGAPLLGWNRYILDVHGLGCTVDWKSKDANDSSFVLFLFLGCLVVPLGVI
AHCYGHILYSIRMLRCVEDLQTIQVIKILKYEKKLAKMCFLMIFTFLVCWMPYIVICF
LVVNGHGHLVTPTISIVSYLFAKSNTVYNPVIYVFMIRKFRRSLLQLLCLRLLRCQRP
AKDLPAAGSEMQIRPIVMSQKDGDRPKKKVTFNSSSIIFIITSDESLSVDDSDKTIGV
QSLMLIQVRPL"

BASE COUNT 522 a 516 c 480 g 592 t ORIGIN

В	query Match 89.8%; Score 1380.8; DB 9; Length 2110; est Local Similarity 98.3%; Pred. No. 2.6e-214; atches 1421; Conservative 0; Mismatches 12; Indels 13; Gaps 2;
Qy Db	105 cgagccccgccgcaagctgagcgcctccgcccgccaggcgcggcgcgggccatgta 164
Qу	165 ctcggggaaccgcagcggcgccacggctactgggacggcgggggggg
Db	61 CTCGGGGAACCGCAGCGGCCACGGCTACTGGGACGGCGGGGGCCGCGGGCGCTGA 120
Qу	225 ggggccggcggcgggggacactgagccccgcgccctcttcagccccggcacctacga 284
Db	121 GGGGCCGGCGGCGGGCACACTGAGCCCCGCGCCCCTCTTCAGCCCCCGGCACCTACGA 180
Qy	285 gcgcctggcgctgctgctgggctccattgggctgctgggcgtcggcaacaacctgctggt 344
Db Qy	181 GCGCCTGGCGCTGCTGGGCTCCATTGGGCTGCTGGGCGTCGGCAACAACCTGCTGGT 240
Db	345 gctcgtcctctactacaagttccagcggctccgcactccactcact
Qу	405 catcagcetcagegacetgetggtgtecetetteggggtcacetttacettegtgtacet
Db	
Qy	465 cctgaggaacqqctqqqtqtqqqacaccqtqqqctqcqtqtqqqacqqqtttaqqqac
Db	
Qy	525 cctcttcgggattgtttccattgccaccctaaccgtgctggcctatgaacgttacattcg 584
Db	421 CCTCTTCGGGATTGTTTCCATTGCCACCCTAACCGTGCTGGCCTATGAACGTTACATTCG 480
Qу	585 cgtggtccatgccagagtgatcaatttttcctgggcctggagggccattacctacatctg 644
Db Qy	461 CG1GG1CCATGCCAGAGTGATCAATTTTTCCTGGGCCTGGAGGGCCATTACCTACATCTG 540
Db	645 gctctactcactggcgtgggcaggagcacctctcctgggatggaacaggtacatcctgga 704
Qу	705 cgtacacggactaggctgcactgtggactggaaatccaaggatgccaacgattgctactt
Db	
Qy	765 tgtgettttettatttettggetgeetggtggtgeetgggtgtestaggegettggta
Db	
Qy	825 tggccatattctatattccattcgaatgcttcgttgtgtgggaagatcttcagacaattca 884
Db	721 IGGCCATATTCTATATTCCATTCGAATGCTTCGTTGTGTGGAAGATCTTCAGACAATTCA 780
Qу	885 agtgatcaagattttaaaatatgaaaagaaactggccaaaatgtgctttttaatgatatt 944
Db	781 AGTGATCAAGATTTTAAAATATGAAAAGAAACTGGCCAAAATGTGCTTTTTAATGATATT 840

Qу	945	caccttcctggtctgttggatgccttatatcgtgatctgcttcttggtggttaatggtca	1004
Db	841	CACCTTCCTGGTCTGTTGGATGCCTTATATCGTGATCTGCTTCTTGGTGGTTAATGGTCA	900
Qу	1005	tggtcacctggtcactccaacaatatctattgtttcgtacctctttgctaaatcgaacac	1064
Db	901	TGGTCACCTGGTCACTCCAACAATATCTATTGTTTCGTACCTCTTTGCTAAATCGAACAC	960
Qу	1065	tgtatacaatccagtgatttatgtcttcatgatcagaaagtttcgaagatcccttttgca	1124
Db	961	TGTATACAATCCAGTGATTATGTCTTCATGATCAGAAAGTTTCGAAGATCCCTTTTGCA	1020
Qу	1125	gettetgtgcctccgactgctgaggtgccagaggcctgctaaagacctaccagcagctgg	1184
Db	1021	GCTTCTGTGCCTCCGACTGCTGAGGTGCCAGAGGCCTGCTAAAGACCTACCAGCAGCTGG	1080
Qy	1185	aagtgaaatgcagatcagacccattgtgatgtcacagaaagatggggacaggccaaagaa	1244
Db	1081	AAGTGAAATGCAGATCAGACCCATTGTGATGTCACAGAAAGATGGGGGACAGGCCAAAGAA	1140
Qу	1245	aaaagtgactttcaactcttcttccatcatttttatcatcaccagtgatgaatcactgtc	1304
Db	1141	AAAAGTGACTTTCAACTCTTCCATCATTTTTATCATCACCAGTGATGAATCACTGTC	1200
Qу	1305	agttgacgacagcgacaaaaccaatgggtccaaagttgatgtaatccaagttcgtcc	1361
Db	1201	AGTTGACGACAGCGACAAAACCATTGGGGTCCAAAGTTTGATGTTAATCCAAGTTCGTCC	1260
Qy	1362	tttgtaggaatgaagaatggcaacgaaagatggggccttaaattggatgccacttttgga	1421
Db	1261	TTTGTAGGAATGAAGGATGGCAACGAAAGGTGGGGCCTTAAATTGGATGCCACTTTTGGA	1320
Qу	1422	ctttcatcataagaagtgtctggaatacccgttctatgtaatatcaacag	1471
Db	1321	CTTTCATCATCCTCCTGAAGAAGAAGTGTCTGGAATACCCGTTCTATGTAATATCAACAG	1380
Qy	1472	aaccttgtggtccagcaggaaatccgaattgcccatatgctcttgggcctcaggaagagg	1531
Db	1381		1440
Qу	1532	ttgaac 1537	
Db	1441	TTGAAC 1446	

SUMMARIES

						COLLINATED	
			*				
Res	ult		Query				
	No.	Score	Match	Length	DB	.ID	Description
	1	1537	100.0	1537	21	AAA38861	Human G-protein co
	2	1459	94.9	2144	21	AAA73212	Human 17723 recept
	3	1373.2	89.3	2037	22	AAD19720	Dendritic cell (DC
	4	1028	66.9	1697	22	AAF33051	Human secreted pro
	5	687.8	44.7	1267	21	AAZ34604	Human receptor mol
	6	644.4	41.9	1763	21	AAC69518	Human secreted pro
	7	437	28.4	619	22	AAD19721	Dendritic cell (DC
C	8	435	28.3	12291	22	AAK79265	Human immune/haema
C	9	424	27.6	5024	24	AAS94874	Human DNA sequence
C	10	422.4	27.5	5000	19	AAV20609	Human kynurenine-3
	11	400.4	26.1	449	20	AAZ42057	Human endometrium

AAA38861 ID AAA3 XX AAA38861 standard; cDNA; 1537 BP.

AC XX AAA38861;

```
DΤ
     31-AUG-2000 (first entry)
 XX
 DE
     Human G-protein coupled receptor, HG51, coding sequence.
 XX
     Human; G-protein coupled receptor; HG51; signal transduction;
 KW
 KW
     rhodopsin receptor; obesity; type II diabetes;
     inflammatory bowel disease; constipation; diarrhoea; gene therapy; ss. .
 KW
 \mathbf{x}\mathbf{x}
 os
     Homo sapiens.
 XX
 FH
     Key
                  Location/Qualifiers
 FT
     CDS
                  160..1368
 FT
                  /*tag= a
 FT
                  /product= "Human HG51"
 XX
 PN
     WO200031108-A1.
 XX
 PD
     02-JUN-2000.
 XX
 PF
     18-NOV-1999;
                 99WO-US27305.
XX
    24-NOV-1998;
                 98US-0109717.
 XX
 PA
     (MERI ) MERCK & CO INC.
XX
ΡI
    Liu Q, McDonald TP;
ХX
DR
    WPI; 2000-400025/34.
DR
    P-PSDB; AAY98008.
XX
PT
    New DNA encoding human HG51 (a G-protein coupled receptor), useful in
    chromosomal mapping studies for identifying the chromosomal locations
PT
PΤ
    of the HG51 gene(s) -
xx
PS
    Claim 1; Fig 1; 68pp; English.
XX
    G protein-coupled receptors (GPCR) are important in signal transduction
CC
    from the exterior to the interior of cells. Rhodopsin receptors are a
CC
    type of GPCR which comprise a chromophore-binding pocket which is
CC
    covalently linked by a protonated Schiff base to a Lys residue in
CC
    transmembrane domain 7. The present sequence is the coding sequence of
CC
    the human HG51 GPCR and is a member of the rhodopsin receptor family of
CC
CC
    GPCRs. Due to the Lys residue and Schiff base present in HG51, it is
    thought that the HG51 ligand may be a fatty-acid-like molecule. It is
CC
CC
    also believed that agonists and antagonists of HG51 are useful for
    treating various disorders such as obesity, type II diabetes,
CC
CC
    inflammatory bowel disease, constipation or diarrhoea. In addition, the
CC
    present sequence may be used in gene therapy for the above mentioned
CC
XX
SQ
    Sequence 1537 BP; 320 A; 426 C; 421 G; 370 T; 0 other;
 Query Match
                     100.0%; Score 1537; DB 21; Length 1537;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1537; Conservative
                          0; Mismatches
                                         0; Indels
                                                               0:
Oν
      1 ggggccacggggggtgcgccggcgggcgggggcccctcagtgcacaatggccag 60
        Db
      Qу
     Db
     61 agcaggeggeggageceeageeecagtgeggagegegegegageeeegegeaag 120
Qу
    121 ctgagcgcctccgcccggcaggcgcggcggggccatgtactcggggaaccgcagc 180
        Db
    Qу
```

Db	181 ggcggccacggctactgggacggcggggggcggggggggcgggggggg
Qу	
Db	241 999acactgageceegegeceetetteageceeggeacetacgagegeetggegetgetg 300
Qy Db	301 ctgggetecattgggetgetgggegteggeaacaaectgetggtgetegteetetaetae 360
Qу	361 aagtteeageggeteegeacteecacteaceteeteetggteaacateagecteagega. 420
Db	
Qy	421 ctgctggtgtccctcttcggggtcacctttaccttcgtgtcctgcctg
Db	
Qy	481 gtgtgggacaccgtgggctgcgtgtgggacgggtttagcggcagcctcttcgggattgtt 540
Db	481 gtgtgggacaccgtgggctgcgtgtgggacgggtttagcggcagcctcttcgggattgtt 540
Qy Db	541 tccattgccaccctaaccgtgctggcctatgaacgttacattcgcgtggtccatgccaga 600
Qy	541 tecattgccaccetaaccgtgctggcctatgaacgttacattcgcgtggtccatgccaga 600
Db	601 gtgatcaatttttcctgggcctggagggccattacctacatctggctctactcactggcg 660
Qу	661 tgggcaggagcacctctcctgggatggaacaggtacatcctggacgtacacggagtacag
Db	
Qy	721 tgcactgtggactggaaatccaaggatgccaacgattcctcctttgtgcttttctcttat
Db	780 721 tgcactgtggactggaaatccaaggatgccaacgattcctcctttgtgcttttcttattt 780
Qy	781 cttggctgcctggtggtgcccctgggtgtcatagcccattgctatggccatattctatat 840
Db	781 Cttggctgctggtggtgccctgggtgtcatagcccattgctatggccatattctatat 840
Qy Db	841 tccattcgaatgcttcgttgtgtggaagatcttcagacaattcaagtgatcaagatttta 900
Qy	841 tccattcgaatgcttcgttgtgtggaagatcttcagacaattcaagtgatcaagatttta 900 901 aaatatgaaaagaaactggccaaaatgtgctttttaatgatattcaccttcctggtctgt 960
Db	
Qy	961 tggatgccttatatcgtgatctgcttcttggtggttaatggtcatggtcacctggtgatt 1020
Db	
Qу	1021 ccaacaatatctattgtttcgtacctctttgctaaatcgaacactgtatacaatccagtg 1080
Db	1021 Ccaacaatatctattgtttcgtacctctttgctaaatcgaacactgtatacaatccagtg 1080
Qy	1081 atttatgtcttcatgatcagaaagtttcgaagatcccttttgcagcttctgtgcctccga 1140
Db Qy	1001 atttatgtetteatgateagaaagtttegaagateeettttgeagettetgtgeeteega 1140
Db	1141 ctgctgaggtgccagaggcctgctaaagacctaccagcagctggaagtgaaatgcagatc 1200
Qy	1201 agacccattgtgatgtcacagaaagatggggacaggccaaagaaaaagtgagtttgaag 1260
Db	
Qу	1261 tcttcttccatcatttttatcatcaccagtgatgaatcactgtcagttgacgacagcgac 1320

```
Db
      1261 tcttcttccatcatttttatcatcaccagtgatgaatcactgtcagttgacgacagcgac 1320
 Qу
      1321 aaaaccaatgggtccaaagttgatgtaatccaagttcgtcctttgtaggaatgaagaatg 1380
           Db
      1321 aaaaccaatgggtccaaagttgatgtaatccaagttcgtcctttgtaggaatgaagaatg 1380
      1381 gcaacgaaagatggggccttaaattggatgccacttttggactttcatcataagaagtgt 1440
 0ν
           Db
      1381 gcaacgaaagatggggccttaaattggatgccacttttggactttcatcataagaagtgt 1440
 Qу
      1441 ctggaatacccgttctatgtaatatcaacagaaccttgtggtccagcaggaaatccgaat 1500
           Db
      1441 ctggaatacccgttctatgtaatatcaacagaaccttgtggtccagcaggaaatccgaat 1500
      1501 tgcccatatgctcttgggcctcaggaagaggttgaac 1537
 Qу
           !!!!!!!!!!!!!!!!!!!!!!!!
 Db
      1501 tgcccatatgctcttgggcctcaggaagaggttgaac 1537
 RESULT
 AAA73212
     AAA73212 standard; cDNA; 2144 BP.
 XX
 AC
     AAA73212;
 XX
     05-DEC-2000 (first entry)
 DT
 XX
 DE
     Human 17723 receptor protein encoding cDNA SEQ ID NO:2.
 XX
 KW
     Human; 17723 receptor protein; chromosome 1q42-44; diagnosis; vaccine;
 KW
     G-protein coupled receptor; gene therapy; ss.
 XX
 os
     Homo sapiens.
XX
 PN
     WO200043513-A1.
XX
PD
     27-JUL-2000.
XX
PF
     21-JAN-2000; 2000WO-US01592.
XX
PR
     21-JAN-1999;
                   99US-0234923.
XX
     (MILL-) MILLENNIUM PHARM INC.
PA
XX
ΡI
     Glucksmann MA;
XX
     WPI; 2000-476196/41.
DR
DR
     P-PSDB; AAB12827.
XX
PT
     A G-protein-coupled receptor designated 17723 and the nucleic acids
     that encode it, useful for preventing, diagnosing and treating disorder
PT
     associated with inappropriate expression of 17723 receptors -
PT
XX
PS
    Claim 3; Page 72-73; 79pp; English.
XX
CC
    The present sequence encodes the human 17723 receptor protein (I), which
CC
    belongs to the superfamily of G-protein-coupled receptors. (I) and the
CC
    polynucleotide encoding it may be used in the prevention, treatment and
CC
    diagnosis of diseases associated with inappropriate 17723 receptor
    expression. They may also be used to study the expression and function
CC
CC
    of 17723 receptor polypeptides and their role in metabolism. The 17723
    receptor polypeptides may be used as antigens in the production of
CC
CC
    antibodies against 17723 receptors and in assays to identify modulators
    (agonists and antagonists) of 17723 receptor expression and activity.
CC
CC
    The anti-17723 receptor antibodies and 17723 receptor antagonists may be
CC
    used to down regulate 17723 receptor expression and activity. The
CC
    anti-17723 receptor antibodies may also be used as diagnostic agents for
CC
    detecting the presence of 17723 receptor polypeptides in samples
CC
    (e.g. by enzyme linked immunosorbent assay (ELISA)). The 17723 receptor
    protein has been mapped to chromosome 1q42-44.
CC
```

SQ Sequence 2144 BP; 525 A; 531 C; 496 G; 590 T; 2 other;

E	Query Match 94.9%; Score 1459; DB 21; Length 2144; Best Local Similarity 99.9%; Pred. No. 0; Matches 1470; Conservative 0; Mismatches 0; Indels 1; Gaps	1;
Qу	67 eggeggagececagececacecagtgeggagegegegeggageceegeegeaagetgage 126	
Db	1 cggcggagccccag-cccaccagtgcggagcgcgcgcgcgcgcgccgccgccgccgc59	
Qу	127 gcctccgccaggcgcgcggcgcgggccatgtactcggggaaccgcagcggcggc 186	
Db	60 gcctccgccaggcgcggcggcgggccatgtactcggggaaccgcagcggcgc 119	
Qу	187 cacggctactgggacggcgggggccgcgggggcgctgaggggccggcggcggcggggaca 246	
Db	120 cacggctactgggacggcgggggccgcgggggctgagggccggcggcggcggggaca 179	
Qу	247 ctgagccccgcgcccctcttcagccccggcacctacgagcgcctggcgctgctgctgggc 306	
Db	180 ctgagccccgcgcccctcttcagccccggcacctacgagcgcctggcgctgctgctggc 239	
Qу	307 tccattgggctgctgggcgtcggcaacaacctgctggtgctcgtcctctactacaagttc 366	
Db	240 tecattgggetgetgggegteggeaacaacetgetggtgetegtectetactacaagtte 299	
Qу	367 cageggeteegeacteecacteaceteeteetggteaacateageeteagegacetgetg 426	
Db	300 Cayeggeteegeacteccacteacetectggteaacateagecteagegacetgetg 359	
Qу	427 gtgtccctcttcggggtcacctttaccttcgtgtcctgcctg	
Db	300 gtgteectetteggggteacetttacettegtgteetgeetgaggaaeggetgggtgtgg 419	
QΥ	487 gacaccgtgggctgcgtgtgggacgggtttagcggcagcctcttcgggattgtttccatt 546	
Db	420 gacaccgtgggctgcgtgtgggacgggtttagcggcagcctcttcggggattgtttccatt 479	
ДУ	547 gccaccctaaccgtgctggcctatgaacgttacattcgcgtggtccatgccagagtgatc 606	
Db	400 gccaccctaaccgtgctggcctatgaacgttacattcgcgtggtccatgccagagtgatc 539	
Qy Db	607 aatttttcctgggcctggagggccattacctacatctggctctactcactggcgtgggca 666	
Qу	340 dattttttttttgggeetggagggeeattacetacatetggetetactcaetggegtgggca 599	
Db	667 ggagcacctctcctgggatggaacaggtacatcctggacgtacacggactaggctgcact 726	
Qу	600 ggagcacctctcctgggatggaacaggtacatcctggacgtacacggactaggctgcact 659	
Db	727 gtggactggaaatccaaggatgccaacgattcctcctttgtgcttttcttatttcttggc 786	
Qy	660 gtggactggaaatccaaggatgccaacgattcctcctttgtgcttttcttatttcttggc 719 787 tgcctggtggtgcccctgggtgtcatagcccattgctatggccatattctatattccatt 846	
Db	720 tgcctggtggtgccctgggtgtcatagcccattgctattgccatttctatttccatt 779	
Qу	847 cgaatgcttcgttgtgtggaagatcttcagacaattcaagtgatcaagattttaaaatat 906	•
Db	780 cgaatgcttcgttgtgtgaagatcttcagacaattcaagtgatcaagattttaaaatat 906	
Qy	907 gaaaagaaactggccaaaatgtgctttttaatgatattcaccttcctggtctgttggatg	
Db		
Qу	967 ccttatatcgtgatctgcttcttggtggttaatggtcatggtcacctggtcactggatcacc	
Db		

0	1000		
Qу	1027	atatctattgtttcgtacctctttgctaaatcgaacactgtatacaatccagtgatttat /	1086
Db	960	atatctattgtttcgtacctctttgctaaatcgaacactgtatacaatccagtgatttat	1019
Qу	1087	gtcttcatgatcagaaagtttcgaagatcccttttgcagcttctgtgcctccgactgctg	1146
Db	1020		1079
Qу	1147	aggtgccagaggcctgctaaagacctaccagcagctggaagtgaaatgcagatcagaccc	1206
Db	1080	aggtgccagaggcctgctaaagacctaccagcagctggaagtgaaatgcagatcagaccc	1139
Qy	1207	attgtgatgtcacagaaagatggggacaggccaaagaaaaagtgactttcaactcttct	1266
Db	1140	attgtgatgtcacagaaagatggggacaggccaaagaaaaagtgactttcaactcttct	1199
Qу	1267	tccatcatttttatcatcaccagtgatgaatcactgtcagttgacgacagcgacaaaacc	1326
Db	1200		1259
Qy	1327	aatgggtccaaagttgatgtaatccaagttcgtcctttgtaggaatgaagaatggcaacg	1386
Db	1260		1319
Qy	1387		1446
Db	1320		1379
Qу	1447	tacccgttctatgtaatatcaacagaaccttgtggtccagcaggaaatccgaattgccca	1506
Db			
Qy		tatgctcttgggcctcaggaagaggttgaac 1537	
Db	1440		

Result No. 5	Score	Query Match	Length	DB	ID	Description
2 1 3 1 4 5 6 7 8 9 10	122.4 107.6 107.6 78 73 69.2 69.2 67.2 67.2 67.4 65.4	27.5 7.0 7.0 5.1 4.5 4.5 4.4 4.4 4.3 4.3	5000 3016 3016 1105 1410 1420 1420 1293 1776 2140 2140	3 1 2 4 1 3 4 1 1	US-09-147-522-5 US-07-805-123C-1 US-08-033-081B-1 US-08-466-103A-15 US-09-255-368-1 US-08-358-171-1 US-09-090-947-1 US-09-255-368-7 US-08-722-001-29 US-08-334-698-1 US-08-228-932-1 US-08-468-939-1	Sequence 5, Appli Sequence 1, Appli Sequence 1, Appli Sequence 15, Appl Sequence 1, Appli Sequence 1, Appli Sequence 1, Appli Sequence 29, Appl Sequence 1, Appli Sequence 1, Appli Sequence 1, Appli Sequence 1, Appli

Result		Query				
No.	Score	Match	Length	DB	ID	Description
1	660.8	43.0	789	10	BI818538	BI818538 603033059
2	644.8	42.0	770	10	BI260681	BI260681 602968193
3	609.8	39.7	909	10	BE894106	BE894106 601438234
4	580.8	37.8	736	10	BI086726	BI086726 602850078
5	577.2	37.6	835	10	BF970560	BF970560 602274056
6	575	37.4	850	10	BI757207	BI757207 603030709
7	565.4	36.8	748	10	BG252201	BG252201 602365072
8	515.8	33.6	741	10	BG564220	BG564220 602586010
9	467.8	30.4	788	10	BF977798	BF977798 602148633
10	461.8	30.0	784	10	BI758685	BI758685 603024224

14 406.8 26.5 742 10 BI257225 BI257225 6029 15 398.4 25.9 615 10 BF132059 BF132059 BF132059 6018		2976885
--	--	---------

SEQ ID NO: 2

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	. Description
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16	2117 2117 2105 1063 756 664 572 564 459.5 455 451 449 424 420.5 420.5	100.0 100.0 99.4 50.2 35.7 31.4 27.0 26.6 21.7 21.5 21.3 21.2 20.0 19.9	402 402 402 199 , 147 163 879 123 122 349 348 348 348 354 309	21 22 22 21 22 22 20 21 10 17 21 14 21 15	AAB12827 AAY98008 AAE12070 AAB64743 AAY32195 AAE12071 AAU31008 AAY60172 AAB38327 AAP90554 AAR93116 AAY98009 AAR38483 AAY57086 AAR48735 AAW02707	Human 17723 recept Human G-protein co Dendritic cell (DC Human secreted pro Human receptor mol Dendritic cell (DC Novel human secret Human endometrium Human secreted pro Bovine rhodopsin. Rhodopsin. Homo s Human rhodopsin re Rhodopsin protein. Rhodopsin amino ac G-protein coupled G-protein coupled
						5 procein coupled

```
RESULT
 AAB12827
 ID
      AAB12827 standard; Protein; 402 AA.
 XX
 AC
      AAB12827;
 XX
 DΤ
      05-DEC-2000 (first entry)
 XX
 DΕ
      Human 17723 receptor protein SEQ ID NO:1.
 XX
      Human; 17723 receptor protein; chromosome 1q42-44; diagnosis; vaccine;
 KW
 KW
      G-protein coupled receptor; gene therapy.
XX
os
      Homo sapiens.
XX
PN
      WO200043513~A1.
XX
PD
     27-JUL-2000.
XX
PF
     21-JAN-2000; 2000WO-US01592.
XX
PR
     21-JAN-1999; 99US-0234923.
\mathbf{x}\mathbf{x}
PA
      (MILL-) MILLENNIUM PHARM INC.
XX
ΡI
     Glucksmann MA;
XX
     WPI; 2000-476196/41.
DR
DR
     N-PSDB; AAA73212.
\mathbf{x}\mathbf{x}
     A G-protein-coupled receptor designated 17723 and the nucleic acids
PT
     that encode it, useful for preventing, diagnosing and treating disorder
PT
     associated with inappropriate expression of 17723 receptors -
PT
XX
PS
     Claim 1; Page 70-72; 79pp; English.
XX
```

The present sequence is the human 17723 receptor protein (I), which

```
belongs to the superfamily of G-protein-coupled receptors. (I) and the
  CC
      polynucleotide encoding it may be used in the prevention, treatment and
  CC
  CC
      diagnosis of diseases associated with inappropriate 17723 receptor
      expression. They may also be used to study the expression and function
  CC
      of 17723 receptor polypeptides and their role in metabolism. The 17723
  CC
  CC
      receptor polypeptides may be used as antigens in the production of
      antibodies against 17723 receptors and in assays to identify modulators
  CC
      (agonists and antagonists) of 17723 receptor expression and activity.
  CC
      The anti-17723 receptor antibodies and 17723 receptor antagonists may be
  CC
      used to down regulate 17723 receptor expression and activity. The
  CC
      anti-17723 receptor antibodies may also be used as diagnostic agents for
  CC
      detecting the presence of 17723 receptor polypeptides in samples
      (e.g. by enzyme linked immunosorbent assay (ELISA)). The 17723 receptor
  CC
  CC
      protein has been mapped to chromosome 1q42-44.
  XX
 SO
      Sequence
                402 AA;
   Query Match
                         100.0%; Score 2117; DB 21; Length 402;
   Best Local Similarity 100.0%; Pred. No. 6.3e-222;
   Matches 402; Conservative
                               0; Mismatches
                                                0; Indels
        1 MYSGNRSGGHGYWDGGGAAGAEGPAPAGTLSPAPLFSPGTYERLALLLGSIGLLGVGNNL 60
 Qу
          nnnammunnuminnijananjijijijijijijimij
 Db
        1 mysgnrsgghgywdgggaagaegpapagtlspaplfspgtyerlalllgsigllgvgnnl 60
       61 LVLVLYYKFQRLRTPTHLLLVNISLSDLLVSLFGVTFTFVSCLRNGWVWDTVGCVWDGFS 120
 Qу
          յուսուցուսուսությունությունություն անություն
       61 lvlvlyykfqrlrtpthlllvnislsdllvslfgvtftfvsclrngwvwdtvgcvwdgfs 120
 Db
      121 GSLFGIVSIATLTVLAYERYIRVVHARVINFSWAWRAITYIWLYSLAWAGAPLLGWNRYI 180
 Qу
          121 gslfgivsiatltvlayeryirvvharvinfswawraityiwlyslawagapllgwnryi 180
 Db
      181 LDVHGLGCTVDWKSKDANDSSFVLFLFLGCLVVPLGVIAHCYGHILYSIRMLRCVEDLQT 240
 Qу
          Db
      181 ldvhglgctvdwkskdandssfvlflflgclvvplgviahcyghilysirmlrcvedlqt 240
      241 IQVIKILKYEKKLAKMCFLMIFTFLVCWMPYIVICFLVVNGHGHLVTPTISIVSYLFAKS 300
 Qу
          241 iqvikilkyekklakmcflmiftflvcwmpyivicflvvnghghlvtptisivsylfaks 300
 Db
      301 NTVYNPVIYVFMIRKFRRSLLQLLCLRLLRCQRPAKDLPAAGSEMQIRPIVMSQKDGDRP 360
Qy
          អាការប្រជាជាក្រាយប្រើប្រែប្រជាជាក្រាយប្រជាជាក្រាយប្រជាជាក្រា
      301 ntvynpviyvfmirkfrrsllqllclrllrcqrpakdlpaagsemqirpivmsqkdgdrp 360
Db
      361 KKKVTFNSSSIIFIITSDESLSVDDSDKTNGSKVDVIQVRPL 402
Qy
          361 kkkvtfnsssiifiitsdeslsvddsdktngskvdviqvrpl 402
RESULT
AAY98008
TD
    AAY98008 standard; Protein; 402 AA.
XX
AC
    AAY98008;
XX
DТ
    31-AUG-2000 (first entry)
XX
DE
    Human G-protein coupled receptor, HG51.
XX
KW
    Human; G-protein coupled receptor; HG51; signal transduction;
    rhodopsin receptor; obesity; type II diabetes;
KW
    inflammatory bowel disease; constipation; diarrhoea; gene therapy.
KW
XX
OS
    Homo sapiens.
XX
PN
    WO200031108-A1.
ХX
PD
    02-JUN-2000.
XX
```

```
PF
      18-NOV-1999;
                   99WO-US27305.
 XX
 PR
      24-NOV-1998;
                   98US-0109717.
 XX
 PA
      (MERI ) MERCK & CO INC.
 XX
 PΙ
      Liu Q, McDonald TP;
 XX
 DR
      WPI; 2000-400025/34.
 DR
      N-PSDB; AAA38861.
 XX
 PT
      New DNA encoding human HG51 (a G-protein coupled receptor), useful in
      chromosomal mapping studies for identifying the chromosomal locations
 PT
      of the HG51 gene(s) -
 PT
 XX
 PS
     Claim 23; Fig 2; 68pp; English.
 XX
 CC
     G protein-coupled receptors (GPCR) are important in signal transduction
     from the exterior to the interior of cells. Rhodopsin receptors are a
 CC
     type of GPCR which comprise a chromophore-binding pocket which is
 CC
 CC
     covalently linked by a protonated Schiff base to a Lys residue in
     transmembrane domain 7. The present sequence is the human HG51 GPCR and
     is a member of the rhodopsin receptor family of GPCRs. Due to the Lys
 CC
     residue and Schiff base present in HG51, it is thought that the HG51
     ligand may be a fatty-acid-like molecule. It is also believed that
 CC
     agonists and antagonists of HG51 are useful for treating various
 CC
     disorders such as obesity, type II diabetes, inflammatory bowel disease,
 CC
     constipation or diarrhoea. In addition, the coding sequence for the
     present sequence may be used in gene therapy for the above mentioned
 CC
 CC
     disorders.
 XX
 SQ
     Sequence
               402 AA:
  Query Match
                       100.0%; Score 2117; DB 21; Length 402;
  Best Local Similarity 100.0%; Pred. No. 6.3e-222;
  Matches 402; Conservative
                             0; Mismatches
                                             0; Indels
                                                                     0;
        1 MYSGNRSGGHGYWDGGGAAGAEGPAPAGTLSPAPLFSPGTYERLALLLGSIGLLGVGNNL 60
Ov
         {\tt 1~mysgnrsgghgywdgggaagaegpapagtlspaplfspgtyerlalllgsigllgvgnnl~60}\\
0v
       61 LVLVLYYKFQRLRTPTHLLLVNISLSDLLVSLFGVTFTFVSCLRNGWVWDTVGCVWDGFS 120
          61 lvlvlyykfqrlrtpthlllvnislsdllvslfgvtftfvsclrngwvwdtvgcvwdgfs 120
      121 GSLFGIVSIATLTVLAYERYIRVVHARVINFSWAWRAITYIWLYSLAWAGAPLLGWNRYI 180
Qу
          Db
      121 gslfgivsiatltvlayeryirvvharvinfswawraityiwlyslawagapllgwnryi 180
      181 LDVHGLGCTVDWKSKDANDSSFVLFLFLGCLVVPLGVIAHCYGHILYSIRMLRCVEDLQT 240
Qу
          ummanaanimmännimmänniminiimiiliiniimiimiili
Db
     181 ldvhglgctvdwkskdandssfvlflflgclvvplgviahcyghilysirmlrcvedlqt 240
     241 IQVIKILKYEKKLAKMCFLMIFTFLVCWMPYIVICFLVVNGHGHLVTPTISIVSYLFAKS 300
Qу
         Db
     241 iqvikilkyekklakmcflmiftflvcwmpyivicflvvnghghlvtptisivsylfaks 300
     301 NTVYNPVIYVFMIRKFRRSLLQLLCLRLLRCQRPAKDLPAAGSEMQIRPIVMSQKDGDRP 360
Qу
         301 ntvynpviyvfmirkfrrsllqllclrllrcqrpakdlpaagsemqirpivmsqkdgdrp 360
Db
     361 KKKVTFNSSSIIFIITSDESLSVDDSDKTNGSKVDVIQVRPL 402
Qу
         361 kkkvtfnsssiifiitsdeslsvddsdktngskvdviqvrpl 402
Db
RESULT
```

AAE12070

ID AAE12070 standard; Protein; 402 AA.

```
AC
       AAE12070:
  XX
  DT
       18-DEC-2001 (first entry)
  XX
  DE
       Dendritic cell (DC) DCEPR protein.
  XX
  KW
       Dendritic cell; DC; DCEPR protein; gene therapy; dermatological; vaccine;
       atopic dermatitis; autoimmune disease; inflammatory skin disease; cancer;
  KW
       immunosuppressive; AIDS; Acquired immune deficiency syndrome; cytostatic;
  KW
  KW
       chromosomal identification; pharmaceutical; hypersensitivity; virucide;
  KW
       transplant rejection; chronic inflammatory disease; anti-HIV.
 XX
 OS
       Unidentified.
 XX
 PN
       WO200172773-A2.
 XX
 PD
       04-OCT-2001.
 XX
 PF
      28-MAR-2001; 2001WO-EP03542.
 XX
 PR
      29-MAR-2000; 2000US-192934P.
 PR
      18-MAY-2000; 2000US-205020P.
      18-MAY-2000; 2000US-205026P.
 PR
 PR
      19-MAY-2000; 2000US-205767P.
      19-MAY-2000; 2000US-205769P.
 PR
 XX
       (NOVS ) NOVARTIS AG.
 PA
 PA
      (NOVS ) NOVARTIS-ERFINDUNGEN VERW GES MBH.
 XX
 PΙ
      Werner G, Phares W, Jaritz M, Lapp H, Kalthoff FS;
 XX
      WPI; 2001-616466/71.
 DR
 DR
      N-PSDB; AAD19720.
 XX
 PT
      New polypeptides for screening therapeutic agonists and antagonists
 PT
      comprise dendritic cell polypeptides -
 XX
 PS
      Claim 2; Page 42; 52pp; English.
 XX
      The invention relates to dendritic cell (DC) proteins and their
 CC
 CC
      corresponding DNA molecules. A pharmaceutical composition comprising
      agonist and antagonist of DC proteins are useful for treating abnormal
 CC
 CC
      conditions related to both an excess of and insufficient level of
      expression of DC gene, or related to both an excess of and insufficient
CC
     activity of DC protein. Soluble form of DC proteins are used as an active
CC
CC
      ingredient in combination with pharmaceutical acceptable carriers.
     DC genes and proteins are useful for treating chronic inflammatory
CC
     diseases, autoimmune diseases, transplant rejection crisis, including
CC
     inflammatory skin diseases such as contact hypersensitivity, atopic
CC
     dermatitis or virally-induced immune suppression such as AIDS and cancer.
CC
     DC protein is useful for inducing immunological response in a mammal, and
CC
     as immunogen to produce antibodies immunospecific for the polypeptide.
CC
     DC gene is useful in gene therapy. DC gene is also useful as a diagnostic
CC
     reagent, and for chromosomal identification. The present sequence is
CC
     dendritic cell (DC) DCEPR protein which is found to belong to the family
CC
CC
     of G-protein coupled receptor protein.
XX
so
     Sequence
                402 AA;
                           99.4%; Score 2105; DB 22;
                                                        Length 402;
  Best Local Similarity
                           99.5%; Pred. No. 1.3e-220;
  Matches 400; Conservative
                                  1; Mismatches
                                                    1;
                                                        Indels
                                                                               0;
        1 MYSGNRSGGHGYWDGGGAAGAEGPAPAGTLSPAPLFSPGTYERLALLLGSIGLLGVGNNL 60
Qy
          MINITERINALITATE: DELIBERATION DE LA MINITERIORI DE LA MINITERIORI DE LA MINITERIORI DE LA MINITERIORI DE LA M
Db
        {\tt 1~mysgnrsgghgywdgggaagakgpapagtlspaplfspgtyerlalllgsigllgygnnl~60}\\
       61 LVLVLYYKFQRLRTPTHLLLVNISLSDLLVSLFGVTFTFVSCLRNGWVWDTVGCVWDGFS 120
Qу
          អាយាម្យាយអាជីវិយល់បានប្រជាវិយលីលេខ បានប្រជាវិយាយលើ
Db
       61 lvlvlyykfqrlrtpthlllvnislsdllvslfgvtftfvsclrngwvwdtvgcvwdgfs 120
```

```
121 GSLFGIVSIATLTVLAYERYIRVVHARVINFSWAWRAITYIWLYSLAWAGAPLLGWNRYI 180
  Qу
           121 gslfgivsiatltvlayeryirvvharvinfswawraityiwlyslawagapllgwnryi 180
 Db
       181 LDVHGLGCTVDWKSKDANDSSFVLFLFLGCLVVPLGVIAHCYGHILYSIRMLRCVEDLQT 240
 Qу
           181 ldvhglgctvdwkskdandssfvlflflgclvvplgviahcyghilysirmlrcvedlqt 240
 Db
       241 IQVIKILKYEKKLAKMCFLMIFTFLVCWMPYIVICFLVVNGHGHLVTPTISIVSYLFAKS 300
 Qу
           Db
       241 iqvikilkyekklakmcflmiftflvcwmpyivicflvvnghghlvtptisivsylfaks 300
       301 NTVYNPVIYVFMIRKFRRSLLQLLCLRLLRCQRPAKDLPAAGSEMQIRPIVMSQKDGDRP 360
 Qу
           Db
       301 ntvynpviyvfmirkfrrsllqllclrllrcqrpakdlpaagsemqirpivmsqkdgdrp 360
 Qу
       361 KKKVTFNSSSIIFIITSDESLSVDDSDKTNGSKVDVIQVRPL 402
          Db
       361 kkkvtfnsssiifigtsdeslsvddsdktngskvdviqvrpl 402
 RESULT
 AAB64743
 ID
     AAB64743 standard; Protein; 199 AA.
 AC
     AAB64743;
 XX
 DT
     23-MAR-2001 (first entry)
 DE
     Human secreted protein sequence encoded by gene 15 SEQ ID NO:137.
 XX
     Human; secreted protein; diagnosis; cytostatic; antirheumatic;
 KW
     antiarthritic; dermalogical; cardiant; antiinflammatory; anti-ulcer;
 KW
     gastrointestinal; solid tumour; rheumatoid arthritis; psoriasis;
 KW
     diabetic retinopathy; myocardial angiogenesis; Crohn's disease;
 ĸw
     ulcer.
 XX
 os
     Homo sapiens.
 XX
 PN
     WO200077237-A1.
 XX
 PD
     21-DEC-2000.
XX
 PF
     01-JUN-2000; 2000WO-US14928.
XX
 PR
     11-JUN-1999;
                  99US-0138633.
XX
PA
     (HUMA-) HUMAN GENOME SCI INC.
PΑ
     (ROSE/) ROSEN C A.
XX
PΤ
     Rosen CA, Ruben SM, Komatsoulis GA;
XX
DR
     WPI; 2001-071280/08.
XX
PT
     Nucleic acids encoding 49 human secreted polypeptides, useful for
     preventing, diagnosing and/or treating diseases such as tumors,
PT
     rheumatoid arthritis, psoriasis and diabetic retinopathy -
PT
XX
PS
    Disclosure; Page 503; 520pp; English.
XX
CC
    The polynucleotide sequences given in AAF33037 to AAF33085 encode the
CC
    human secreted proteins given in AAB64666 to AAB64714. AAB64715 to
    AAB64771 represent human secreted polypeptide sequences and proteins
CC
    homologous to them, which are given in the exemplification of the present
CC
    invention. Human secreted proteins have activities based on the tissues
CC
CC
    and cells the genes are expressed in. Examples of activities include:
    cytostatic; antirheumatic; antiarthritic; dermalogical; cardiant;
CC
CC
    antiinflammatory; gastrointestinal; and anti-ulcer. The polynucleotides
    and polypeptides can be used in the prevention, treatment and diagnosis
CC
    of diseases associated with inappropriate polypeptide expression.
```

Disorders that may be treated or prevented include solid tumours, CC CC rheumatoid arthritis, psoriasis, diabetic retinopathy, myocardial CC angiogenesis, Crohn's disease and ulcers. The polynucleotides and their complementary sequences may also be used as DNA probes in diagnostic CC assays (e.g. polymerase chain reactions (PCR)) to detect and quantitate CC the presence of similar nucleic acid sequences in samples, and therefore which patients may be in need of restorative therapy. The polypeptides CC may also be used as antigens in the production of antibodies against the polypeptide and in assays to identify modulators (agonists and CC antagonists) of polypeptide expression and activity. The anti-polypeptide CC antibodies and antagonists may also be used to down regulate expression CC and activity. AAF33028 to AAF33036 and AAB64665 represent sequences used in the exemplification of the present invention. CC SQ Sequence 199 AA; Query Match

Query Match 50.2%; Score 1063; DB 22; Length 199; Best Local Similarity 100.0%; Pred. No. 2e-107; Matches 199; Conservative 0; Mismatches 0; Indels 0; Gaps 0; 118 GFSGSLFGIVSIATLTVLAYERYIRVVHARVINFSWAWRAITYIWLYSLAWAGAPLLGWN 177 Db 1 gfsgslfgivsiatltvlayeryirvvharvinfswawraityiwlyslawagapllgwn 60 178 RYILDVHGLGCTVDWKSKDANDSSFVLFLFLGCLVVPLGVIAHCYGHILYSIRMLRCVED 237 Qу <u>រូបអាការអាចក្រុមប្រជាពលរបស់</u>ពេលការប្រាស់អាវិការប្រែប្រែប្រែក ប្រែប្រែក្រុមប្រែក្រុមប្រែក្រុមប្រែក្រុមប្រែក្រុមប្រ Db 61 ryildvhglgctvdwkskdandssfvlflflgclvvplgviahcyghilysirmlrcved 120 238 LQTIQVIKILKYEKKLAKMCFLMIFTFLVCWMPYIVICFLVVNGHGHLVTPTISIVSYLF 297 121 lqtiqvikilkyekklakmcflmiftflvcwmpyivicflvvnghghlvtptisivsylf 180 Db Qу 298 AKSNTVYNPVIYVFMIRKF 316 11111111111111111111 Db 181 aksntvynpviyvfmirkf 199

Result	G	% Query				
No.	Score	Match	Length	DB	ID	Description
1 2 3 4 5 6 7 8 9 10	451 444 420.5 420.5 354.5 354.5 341.5 341.5 338.5	21.3 21.0 19.9 19.9 16.7 16.1 16.1	348 348 309 309 297 297 305 305 297	2 4 1 5 1 5 1 5	PCT-US93-08528-56 US-08-118-270-58 PCT-US93-08528-58 US-08-118-270-59 PCT-US93-08528-59 US-08-118-270-57 PCT-US93-08528-57	Sequence 8, Appli Sequence 11, Appl Sequence 56, Appl Sequence 56, Appl Sequence 58, Appl Sequence 58, Appl Sequence 59, Appl Sequence 59, Appl Sequence 57, Appl Sequence 57, Appl
12 13 14 15	309 309 309 304 304	14.6 14.6 14.6 14.4 14.4	391 391 391 391 391	1 1 1 1	US-07-816-283-2 US-08-417-103-2 US-08-417-103-14 US-07-816-283-4 US-08-417-103-4	Sequence 2, Appli Sequence 2, Appli Sequence 14, Appli Sequence 4, Appli Sequence 4, Appli

Result No.	Score	Query Match	Length	DB	ID	Description
1 2 3 4 5 6	477.5 475 464 458 456.5 455	22.6 22.4 21.9 21.6 21.6 21.5	349 351 352 348 348 348	1 1 2 1 1	JC5490 A55962 I50081 OOBO JC4267 S23398	opsin, pineal glan opsin, pineal glan rhodopsin - green rhodopsin - bovine opsin - rabbit rhodopsin - Chines
7	452.5	21.4	351	2	S29152	rhodopsin - chicke

```
348 1 OOHU
                                                              rhodopsin - human
       9
                           354 1 S27231
348 1 A23665
             451
                   21.3
                                                              rhodopsin - northe
      10
                   21.3
                                                              opsin - mouse
      11
             448
                   21.2
                           354 1 I51200
                                                              rhodopsin - Africa
  Result
                  Query
     No.
           Score Match Length DB ID
                                                              Description
            2117 100.0
       1
                           402 1 OPN3_HUMAN
                                                              Q9h1y3 homo sapien
       2
            1862 88.0
                         400 1 OPN3_MOUSE
                                                              Q9wuk7 mus musculu
                           349 1 OPSP_COLLI
351 1 OPSP_CHICK
       3
           477.5
                   22.6
                                                           P51476 columba liv
            475
                   22.4
                                                            P51475 gallus gall
       5
             468
                           352 1 OPSD_ALLMI
                   22.1
                                                             P52202 alligator m
       6
           466.5
                   22.0
                           444 1 OPSP PETMA
                                                             042490 petromyzon
                           352 1 OPSD_ANOCA
348 1 OPSD_BOVIN
             464
                   21.9
                                                            P41591 anolis caro
       8
             458
                   21.6
                                                            P02699 bos taurus
          456.5
      9
                   21.6
                           348 1 OPSD_RABIT
                                                             P49912 oryctolagus
      10
            455
455
                  21.5
                                                            P28681 cricetulus
                           348 1 OPSD CRIGR
      11
                           348 1 OPSD_MACFA
354 1 OPSD_RANCA
                   21.5
                                                            Q28886 macaca fasc
             455 21.5
                                                             P51470 rana catesb
 RESULT
 OPN3 HUMAN
 ID
      OPN3_HUMAN
                     STANDARD;
                                     PRT;
                                            402 AA.
      Q9H1Y3; Q9Y344;
      16-OCT-2001 (Rel. 40, Created)
 DT
      16-OCT-2001 (Rel. 40, Last sequence update)
 DΤ
      01-MAR-2002 (Rel. 41, Last annotation update)
      Opsin 3 (Encephalopsin) (Panopsin).
 DE
      OPN3 OR ECPN.
 os
      Homo sapiens (Human).
      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC
      NCBI_TaxID=9606;
 RN
      [1]
      SEQUENCE FROM N.A.
 RP
      MEDLINE=99252448; PubMed=10234000;
 RX
      Blackshaw S., Snyder S.H.;
 RT
      "Encephalopsin: a novel mammalian extraretinal opsin discretely
      localized in the brain.";
 RT
 RL
      J. Neurosci. 19:3681-3690(1999).
 RN
      [2]
 RP
      SEQUENCE FROM N.A.
RX
      MEDLINE=21295039; PubMed=11401433;
      Halford S., Freedman M.S., Bellingham J., Inglis S.L.,
      Poopalasundaram S., Soni B.G., Foster R.G., Hunt D.M.;
RA
      "Characterization of a novel human opsin gene with wide tissue
RT
RT
     expression and identification of embedded and flanking genes on
RT
     chromosome 1q43.";
RL
     Genomics 72:203-208(2001).
RN
      [3]
RΡ
     SEQUENCE FROM N.A.
     Parker A.;
RI.
     Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.
     -!- FUNCTION: May play a role in encephalic photoreception.
CC
     -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC
     -!- TISSUE SPECIFICITY: Strongly expressed in brain. Highly expressed
CÇ
CC
         in the preoptic area and paraventricular nucleus of the
         hypothalamus. Shows highly patterned expression in other regions
CC
CC
         of the brain, being enriched in selected regions of the cerebral
CC
         cortex, cerebellar Purkinje cells, a subset of striatal neurons,
CC
         selected thalamic nuclei, and a subset of interneurons in the
CC
         ventral horn of the spinal cord.
     -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC
         OPSIN SUBFAMILY.
CC
CC
```

8

451

21.3

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

```
CC
      the European Bioinformatics Institute. There are no restrictions on its
 CC
      use by non-profit institutions as long as its content is in no way
 CC
      modified and this statement is not removed. Usage by and for commercial
      entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC
      or send an email to license@isb-sib.ch).
 CC
 CC
      EMBL; AF140242; AAD32671.1; -.
 DR
      EMBL; AF303588; AAK37447.1; -.
 DR
      EMBL; AL133390; CAC19785.1; -.
 DR
      InterPro; IPR000276; GPCR_Rhodpsn.
 DR
      Pfam; PF00001; 7tm 1; 1.
      PRINTS; PR00237; GPCRRHODOPSN.
 DR
      PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
 DR
      PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
 DR
      PROSITE; PS00238; OPSIN; 1.
     Photoreceptor; Retinal protein; Transmembrane; Lipoprotein; Palmitate;
 KW
     G-protein coupled receptor.
     DOMAIN
                 1
                       40
                               EXTRACELLULAR (POTENTIAL).
 FT
     TRANSMEM
                 41
                       65
                               1 (POTENTIAL).
 FT
     DOMAIN
                 66
                      77
                               CYTOPLASMIC (POTENTIAL).
 FТ
     TRANSMEM
                 78
                      102
                               2 (POTENTIAL).
 FT
     DOMATN
                103
                      117
                               EXTRACELLULAR (POTENTIAL).
 FT
     TRANSMEM
                118
                      137
                               3 (POTENTIAL).
 FT
     DOMAIN
                138
                      153
                              CYTOPLASMIC (POTENTIAL).
 FT
     TRANSMEM
                154
                      177
                               4 (POTENTIAL).
 FT
     DOMAIN
                178
                      201
                              EXTRACELLULAR (POTENTIAL).
 FT
     TRANSMEM
                202
                      229
                              5 (POTENTIAL).
 FT
     DOMAIN
                230
                      255
                              CYTOPLASMIC (POTENTIAL).
 FT
     TRANSMEM
                256
                      279
                              6 (POTENTIAL).
 FT
     DOMAIN
                280
                      287
                              EXTRACELLULAR (POTENTIAL).
 FT
     TRANSMEM
               288
                     312
                              7 (POTENTIAL).
 FT
     DOMAIN
               313
                     402
                              CYTOPLASMIC (POTENTIAL).
 FT
     DISULFID
               114
                      188 .
                              BY SIMILARITY.
 FT
     BINDING
               299
                     299
                              RETINAL CHROMOPHORE.
 FT
    LIPID
               325
                      325
                              PALMITATE (BY SIMILARITY).
FT
     CARBOHYD
                5
                      5
                              N-LINKED (GLCNAC. . .) (POTENTIAL).
FT
     CARBOHYD
                              N-LINKED (GLCNAC. . .) (POTENTIAL).
               198
                     198
FT
     CONFLICT
               390
                    396
                              NGSKVDV -> IGVQSLML (IN REF. 1).
     SEQUENCE
              402 AA; 44873 MW; 370F64C19F834A71 CRC64;
  Query Match
                       100.0%; Score 2117; DB 1; Length 402;
  Best Local Similarity 100.0%; Pred. No. 1e-136;
  Matches 402; Conservative
                             0; Mismatches
                                             0; Indels
                                                         0; Gaps
                                                                    0;
       {\tt 1~MYSGNRSGGHGYWDGGGAAGAEGPAPAGTLSPAPLFSPGTYERLALLLGSIGLLGVGNNL~60}\\
Ov
         Db
       1 MYSGNRSGGHGYWDGGGAAGAEGPAPAGTLSPAPLFSPGTYERLALLLGSIGLLGVGNNL 60
      61 LVLVLYYKFQRLRTPTHLLLVNISLSDLLVSLFGVTFTFVSCLRNGWVWDTVGCVWDGFS 120
Qу
         Db
      61 LVLVLYYKFQRLRTPTHLLLVNISLSDLLVSLFGVTFTFVSCLRNGWVWDTVGCVWDGFS 120
     121 GSLFGIVSIATLTVLAYERYIRVVHARVINFSWAWRAITYIWLYSLAWAGAPLLGWNRYI 180
Qу
         Db
     121 GSLFGIVSIATLTVLAYERYIRVVHARVINFSWAWRAITYIWLYSLAWAGAPLLGWNRYI 180
     181 LDVHGLGCTVDWKSKDANDSSFVLFLFLGCLVVPLGVIAHCYGHILYSIRMLRCVEDLQT 240
Ov
         Db
     181 LDVHGLGCTVDWKSKDANDSSFVLFLFLGCLVVPLGVIAHCYGHILYSIRMLRCVEDLQT 240
Qу
     241 IQVIKILKYEKKLAKMCFLMIFTFLVCWMPYIVICFLVVNGHGHLVTPTISIVSYLFAKS 300
         Db
     241 IQVIKILKYEKKLAKMCFLMIFTFLVCWMPYIVICFLVVNGHGHLVTPTISIVSYLFAKS 300
     301 NTVYNPVIYVFMIRKFRRSLLQLLCLRLLRCQRPAKDLPAAGSEMQIRPIVMSQKDGDRP 360
Qу
         301 NTVYNPVIYVFMIRKFRRSLLQLLCLRLLRCQRPAKDLPAAGSEMQIRPIVMSQKDGDRP 360
Db
     361 KKKVTFNSSSIIFIITSDESLSVDDSDKTNGSKVDVIQVRPL 402
Qy
        <del>}</del>
```

Result No.	Score	% Query Match	Length	DB	ID	Description
1	500.5	23.6	352	13	Q9W6K3	Q9w6k3 anolis caro
2	491.5	23.2	534	13	057422	O57422 xenopus lae
3	484.5	22.9	346	13	O9PUA9	Q9pua9 bufo japoni
4	480	22.7	357	13	Q91BH2	Q9ibh2 phelsuma ma
5	473.5	22.4	377	13	O9IB88	Q9ib88 brachydanio
6	473	22.3	543	13	Q90YK6	
7	458	21.6	348	6	095KU1	Q90yk6 gallus gall Q95kul felis silve
8	457.5	21.6	351	13	09IA36	
9	455.5	21.5	351	13	09W6S0	Q9ia36 poephila gu
10	455	21.5	363	13	098TH3	Q9w6s0 columba liv
11	453.5	21.4	322	13	057448	Q98th3 cynops pyrr O57448 anas platyr

SEQ ID NO: 1

					SUMMARIES			
		*	,					
Result	_	Query	_					
No.	Score	Match	Length	DB	ID		Description	
1	1470.4	95.7	2533		AF303588		AF303588 Homo sa	
2	1380.8	89.8	2110	9	AF140242		AF140242 Homo sa	pi
3	943.6	61.4	1737	10	AF140241		AF140241 Mus mu	
c 4	515.4	33.5	71872	9	AL133390		AL133390 Human D	
c 5	424	27.6	5024	6	AX281720		AX281720 Sequenc	
c 6	422.4	27.5	5000	6	A68713		A68713 Sequence	
c 7	422.4	27.5	5000	6	AR106624		AR106624 Sequenc	
c 8	422.4	27.5	5000		AF056032		AF056032 Homo sa	
9	400.4				AX013137		AX013137 Sequence	-
10	358	23.3			AX062411		-	
c 11					AX345325		AX062411 Sequence	
12	299.4	19.5	8303	6	AX345324		AX345325 Sequence	
	233.1	15.5	0303	٠	AA343324		AX345324 Sequence	e
LOCUS DEFINIT: ACCESSIO VERSION KEYWORDS SOURCE ORGAN: REFERENC AUTHOR TITLE JOURNAMEDLIN REFERENC	AF303588 LOCUS AF303588 2533 bp mRNA linear PRI 17-APR-20 DEFINITION Homo sapiens panopsin (OPN3) mRNA, complete cds. ACCESSION AF303588 VERSION AF303588.1 GI:13649589 KEYWORDS . SOURCE human. ORGANISM Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. REFERENCE 1 (bases 1 to 2533) AUTHORS Halford, S., Freedman, M.S., Bellingham, J., Inglis, S.L., Poopalasundaram, S., Soni, B.G., Foster, R.G. and Hunt, D.M.							
JOURNA	L Subn	ect Subm	05-SEP-	200	0) Molecular	Genetics, Ins	titute of	
FEATURES	opnt	almO10	gy, 11-	43	Bath Street,	London EC1V 9	EL, UK	
_	rce			Qua	lifiers			
sou	rce		.2533					
					omo sapiens"			
			_		xon:9606"			
		/c.	hromoso	me=	"1"			
		-	ap="1q4	3"				
gen	e	1.	.2533					
		/g	ene="OF	" ЕИ				
CDS		10	3131	1				
		/ge	ene="OP	'' 8N				
		/no	ote="mu	lti	ple tissue ops	sin"		
			odon st					
					nopsin"			
		/p:	rotein	id=	"AAK37447.1"			
		/d1	xref-	"GT	:13649590"			
		/t:	ranglat	ion	- "MVCCNDCCCtics	WIDCCC A A CA DCDA	D1077 071 71	
		Δ1.1	LGSTGI.	1.011	= MISGNRSGGHG;	WDGGGAAGAEGPA	APAGTLSPAPLFSPGTY	ERL
		ום.ז	ACMAMDA TOTAL	70 V (WDGEGGGI BOTTI WDGEGGGI BOTTI	▗▗▗▗▗▗▗▗ ▗▗▃▗▗▗ ▗ ▞▞▜▜▗▐▗▐▜▜▜▜▜	SLSDLLVSLFGVTFTF	vsc
		TWT	YSIAWA	ים בי	T.CMNDALT.DAMATA	CCALMINGADS	.RVVHARVINFSWAWRA: DSSFVLFLFLGCLVVPLO	TTY
		ZHC	YCHTLV	CIDI	ALDCALDA CULCALA	GCT VDWKSKDANE	SSEVLELELGCLVVPL	σVΙ
		אונן דיי	MChcni.	ハルロロロイン	LIGINGAL DANCE.	. VT PK TEKKTYKUC	FLMIFTFLVCWMPYIV	ICF
		カヤT	חשממטער מיני	OMO:	LISTANCONSCE	VINPVLYVFMIRK	FRRSLLQLLCLRLLRC	QRP
		AVT	VUEAAGS:	r" ⊏MÜ]	LYLINWSÖKDGDKI	KKKVTFNSSSIIF	TITSDESLSVDDSDKT	NGS
		KVL	OVIQVRP	ь"				

В	uery Match 95.7%; Score 1470.4; DB 9; Length 2533; est Local Similarity 99.6%; Pred. No. 7.8e-229; atches 1474; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
Qу	58 cagagcaggcggagccccagcccacccagtgcggagcgcgcgc
Db	1 CGGACGCGTGGGCGCGAGCCCCAGCCCAGTGCGGAGCCCGCGAGCCCCGCCGC 60
Qу	118 aagetgagegeeteegeeegeeaggegeeggeegggeeatgtaeteggggaacege 177
Db	61 AAGCTGAGCGCCCGCCCAGGCGCGCGCGCGCCGGCCATGTACTCGGGGAACCGC 120
Qу	178 agcggcggccacggctactgggacggcggcggggccgcggggcgctgaggggccggcggcgg 237
Db	121 AGCGGCGGCCACGGCTACTGGGACGGCGGCGGCGGCGCGCGC
Qу	238 geggggacactgageeeegegeeetetteageeeeggcacetaegagegeetggegetg 297
Db	181 GCGGGGACACTGAGCCCCGCGCCCCTCTTCAGCCCCGGCACCTACGAGCGCCTGGCGCTG 240
ДУ	298 ctgctgggctccattgggctgctgggcgtcggcaacaacctgctggtgctcgtcctctac 357
Db	241 CTGCTGGGCTCCATTGGGCTGCTGGGCGACAACCTGCTGGTGCTCGTCCTCTAC 300
Qy Db	358 tacaagttccagcggctccgcactcccactcactcctcctggtcaacatcagcctcagc 417
Qy	418 gacctgctggtgtccctcttcggggtcacctttaccttcgtgtcctgcctg
Db	
Qу	478 tgggtgtgggacaccgtgggctgcgtgtgggacgggtttagcggcagcctcttcgggatt 537
Db	
Qy	538 gtttccattgccaccctaaccgtgctggcctatgaacgttacattcgcgtggtccatgcc 597
Db	
Qy	598 agagtgatcaatttttcctgggcctggagggccattacctacatctggctctactcactg 657
Db	541 AGAGTGATCAATTTTTCCTGGGCCTGGAGGGCCATTACCTACATCTGGCTCTACTCACTG 600
Qу	658 gcgtgggcaggagcaccteteetgggatggaacaggtacateetggacgtacaeggaeta 717
Db	601 GCGTGGGCAGGAGCACCTCTCCTGGGATGGAACAGGTACATCCTGGACGTACACGGACTA 660
Qy	718 ggctgcactgtggactggaaatccaaggatgccaacgattcctcctttgtgcttttctta 777
Db	661 GGCTGCACTGTGGACTGGAAATCCAAGGATGCCAACGATTCCTCCTTTGTGCTTTTCTTA 720
Qy Db	778 tttcttggctgctggtggtgcccttgggtgtcatagcccattgctattggccatattcta 837
Qy	721 TTTCTTGGCTGCCTGGTGCCCCTGGGTGTCATAGCCCATTGCTATGGCCATATTCTA 780
Db	838 tattccattcgaatgcttcgttgtgtggaagatcttcagacaattcaagtgatcaagatt 897
Qy	898 ttaaaatatgaaaagaaactqqccaaaatqtqctttttaatgatattcaccttcctqqtc 957
Db	
Qу	958 tgttggatgccttatatcgtgatctgcttcttggtggttaatggtcatggtcacctggtc 1017
Db	901 TGTTGGATGCCTTATATCGTGATCTGCTTCTTGGTGGTTAATGGTCACCTGGTC 960

```
1018 actccaacaatatctattgtttcgtacctctttgctaaatcgaacactgtatacaatcca 1077
 Qу
          961 ACTCCAACAATATCTATTGTTTCGTACCTCTTTGCTAAATCGAACACTGTATACAATCCA 1020
 Db
     1078 gtgatttatgtcttcatgatcagaaagtttcgaagatcccttttgcagcttctgtgcctc 1137
 Qу
     1021 GTGATTTATGTCTTCATGATCAGAAAGTTTCGAAGATCCCTTTTGCAGCTTCTGTGCCTC 1080
 Db
     1138 cgactgctgaggtgccagaggcctgctaaagacctaccagcagctggaagtgaaatgcag 1197
 Qу
         1081 CGACTGCTGAGGTGCCAGAGGCCTGCTAAAGACCTACCAGCAGCTGGAAGTGAAATGCAG 1140
 Db
     {\tt 1198}\ {\tt atcagacccattgtgatgtcacagaaagatggggacaggccaaagaaaaagtgactttc}\ {\tt 1257}
 Qу
         Db
     1141 ATCAGACCCATTGTGATGTCACAGAAAGATGGGGACAGGCCAAAGAAAAAGTGACTTTC 1200
     1258 aactettettecateatttttateateaceagtgatgaateactgteagttgaegaeage 1317
 Qy
         Db
     1201 AACTCTTCTTCCATCATTTTTATCATCACCAGTGATGAATCACTGTCAGTTGACGACAGC 1260
Qy
     1318 gacaaaaccaatgggtccaaagttgatgtaatccaagttcgtcctttgtaggaatgaaga 1377
         Db
     1261 GACAAAACCAATGGGTCCAAAGTTGATGTAATCCAAGTTCGTCCTTTGTAGGAATGAAGA 1320
Qy
     1378 atggcaacgaaagatggggccttaaattggatgccacttttggactttcatcataagaag 1437
         Db
    1321 ATGGCAACGAAAGATGGGGCCTTAAATTGGATGCCACTTTTGGACTTTCATCATAAGAAG 1380
Оy
    1438 tgtctggaatacccgttctatgtaatatcaacagaaccttgtggtccagcaggaaatccg 1497
         1381 TGTCTGGAATACCCGTTCTATGTAATATCAACAGAACCTTGTGGTCCAGCAGGAAATCCG 1440
Qу
    1498 aattgcccatatgctcttgggcctcaggaagaggttgaac 1537
         Db
    1441 AATTGCCCATATGCTCTTGGGCCTCAGGAAGAGGTTGAAC 1480
RESULT 2
AF140242
LOCUS
          AF140242
                               2110 bp
                                        mRNA
                                               linear
                                                     PRI 27-MAY-1999
DEFINITION Homo sapiens encephalopsin mRNA, complete cds.
ACCESSION AF140242
          AF140242.1 GI:4894951
VERSION
KEYWORDS
SOURCE
          human.
  ORGANISM Homo sapiens
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
          1 (bases 1 to 2110)
  AUTHORS
          Blackshaw, S. and Snyder, S.H.
  TITLE
          Encephalopsin: A novel mammalian extraretinal opsin discretely
          localized in the brain
  JOURNAL
          J. Neurosci. 19 (10), 3681-3690 (1999)
 MEDLINE
          99252448
  PUBMED
          10234000
REFERENCE
          2 (bases 1 to 2110)
 AUTHORS
          Blackshaw, S. and Snyder, S.H.
 TITLE
          Direct Submission
 JOURNAL
          Submitted (02-APR-1999) Genetics, Harvard Medical School, 200
          Longwood Ave., Boston, MA 02115, USA
FEATURES
                 Location/Qualifiers
    source
                 1. .2110
                 /organism="Homo sapiens"
                 /db_xref="taxon:9606"
    CDS
                 56. .1267
                 /note="extraretinal opsin"
                 /codon_start=1
                 /product="encephalopsin"
                 /protein_id="AAD32671.1"
                 /db_xref="GI:4894952"
```

/translation="MYSGNRSGGHGYWDGGGAAGAEGPAPAGTLSPAPLFSPGTYERL ALLLGSIGLLGVGNNLLVLVLYYKFQRLRTPTHLLLVNISLSDLLVSLFGVTFTFVSC LRNGWVWDTVGCVWDGFSGSLFGIVSIATLTVLAYERYIRVVHARVINFSWAWRAITY IWLYSLAWAGAPLLGWNRYILDVHGLGCTVDWKSKDANDSSFVLFLFLGCLVVPLGVI AHCYGHILYSIRMLRCVEDLQTIQVIKILKYEKKLAKMCFLMIFTFLVCWMPYIVICF LVVNGHGHLVTPTISIVSYLFAKSNTVYNPVIYVFMIRKFRRSLLQLLCLRLLRCQRP AKDLPAAGSEMQIRPIVMSQKDGDRPKKKVTFNSSSIIFIITSDESLSVDDSDKTIGV QSLMLIQVRPL"

BASE COUNT 522 a 516 c 480 g 592 t ORIGIN

В	uery Match 89.8%; Score 1380.8; DB 9; Length 2110; est Local Similarity 98.3%; Pred. No. 2.6e-214; atches 1421; Conservative 0; Mismatches 12; Indels 13; Gaps 2;	
Qу	105 cgagccccgccaagctgagcgcctccgcccgccaggcgcgcggcgcgggccatgta 164	
Db	1 CGAGCCCCGCCAAGCTGAGCGCCTCCGCCCGCCAGGCGCCGCCGGCCCCGGCCATGTA 60	
Qy Db	165 ctcggggaaccgcagcggcggccacggctactgggacggcgggggggg	
Qy		
Db	225 ggggccggcggcggggacactgagccccgcgccctcttcagccccggcacctacga 284	
Qу	285 gcgcctggcgctgctgctgggctccattgggctgctgggcgtcggcaacaacctgctggt 344	
Db	181 GCGCCTGCCTGCTGCGCTCCATTGGGCTGCTGGGCGTCGCAACCAAC	
Qу	345 gctcgtcctctactacaagttccagcggctccgcactcccactcacctcctcggtcaa 404	
Db	241 GCTCGTCCTACTACAAGTTCCAGCGGCTCCGCACTCCCACTCCTCCTCCTGGTCAA 300	
Qу	405 catcagectcagegacctgctggtgtecetetteggggtcacetttacettegtgteetg 464	
Db		
Qу	465 cctgaggaacggctgggtgtggggacaccgtggggctgcgtgtgggacgggtttagcggcag 524	
Db		
Qу	525 cctcttcgggattgtttccattgccaccctaaccgtgctggcctatgaacgttacattcg 584	
Db		
Qу	585 cgtggtccatgccagagtgatcaatttttcctgggcctggagggccattacctacatctg 644	
Db		
Qy	645 gctctactcactggcgtgggcaggagcacctctcctgggatggaacaggtacatcctgga 704	
Db		
Qy	705 cgtacacggactaggctgcactgtggactggaaatccaaggatgccaacgattcctcctt 764	
Db		
Qу	765 tgtgcttttcttatttcttggctgcctggtggtgcccctgggtgtcatagcccattgcta 824	
Db		
Qу	825 tggccatattctatattccattcgaatgcttcgttgtgtggaagatcttcagacaattca 884	
Db		
Qy	885 agtgatcaagattttaaaatatgaaaagaaactggccaaaatgtgctttttaatgatatt 944	
Db		

Qу	945	caccttcctggtctgttggatgccttatatcgtgatctgcttcttggtggttaatggtca 1004
Db	841	CACCTTCCTGGTCTGTTGGATGCCTTATATCGTGATCTGCTTCTTGGTGGTTAATGGTCA 900
Qy	1005	tggtcacctggtcactccaacaatatctattgtttcgtacctctttgctaaatcgaacac 1064
Db	901	TGGTCACCTGGTCACTCCAACAATATCTATTGTTTCGTACCTCTTTGCTAAATCGAACAC 960
Qу	1065	tgtatacaatccagtgatttatgtcttcatgatcagaaagtttcgaagatcccttttgca 1124
Db	961	TGTATACAATCCAGTGATTTATGTCTTCATGATCAGAAAGTTTCGAAGATCCCTTTTGCA 1020
Qy	1125	gcttctgtgcctccgactgctgaggtgccagaggcctgctaaagacctaccagcagctgg 1184
Db	1021	GCTTCTGTGCCTCCGACTGCTGAGGTGCCAGAGGCCTGCTAAAGACCTACCAGCAGCTGG 1080
Qу	1185	aagtgaaatgcagatcagacccattgtgatgtcacagaaagatggggacaggccaaagaa 1244
Db	1081	AAGTGAAATGCAGATCAGACCCATTGTGATGTCACAGAAAGATGGGGACAGGCCAAAGAA 1140
Qу	1245	aaaagtgactttcaactcttcttccatcatttttatcatcaccagtgatgaatcactgtc 1304
Db	1141	AAAAGTGACTTCAACTCTTCCATCATTTTTATCATCACCAGTGATGAATCACTGTC 1200
Qу	1305	agttgacgacagcgacaaaaccaatgggtccaaagttgatgtaatccaagttcgtcc 1361
Db	1201	AGTTGACGACAGCGACAAAACCATTGGGGTCCAAAGTTTGATGTTAATCCAAGTTCGTCC 1260
Qу	1362	tttgtaggaatgaagaatggcaacgaaagatggggccttaaattggatgccacttttgga 1421
Db	1261	TTTGTAGGAATGAAGGATGGCAACGAAAGGTGGGGCCTTAAATTGGATGCCACTTTTGGA 1320
Qу	1422	ctttcatcataagaagtgtctggaatacccgttctatgtaatatcaacag 1471
Db	1321	CTTTCATCATCCTCCTGAAGAAGAGTGTCTGGAATACCCGTTCTATGTAATATCAACAG 1380
Qy	1472	aaccttgtggtccagcaggaaatccgaattgcccatatgctcttgggcctcaggaagagg 1531
Db	1381	AACCTTGTGGTCCAGCAGGAAATCCGAATTGCCCATATGCTCTTGGGCCTCAGGAAGAGG 1440
Qy	1532	ttgaac 1537
Db	1441	TTGAAC 1446

SUMMARIES

						SUMMARIES	
			₩				
Res	ult		Query				
	No.	Score		Length	DB	ID	Description
	1	1537	100.0	1537	21	AAA38861	Human G-protein co
	2	1459	94.9	2144	21	AAA73212	Human 17723 recept
	3	1373.2	89.3	2037	22	AAD19720	Dendritic cell (DC
	4	1028	66.9	1697	22	AAF33051	• -
	5	687.8	44.7	1267	21	AAZ34604	Human secreted pro
	6	644.4	41.9	1763	21	AAC69518	Human receptor mol
	7	437	28.4	619	22	AAD19721	Human secreted pro
С	8	435	28.3	12291	22	AAK79265	Dendritic cell (DC
С	9	424	27.6	5024	24	AAS94874	Human immune/haema
c	_						Human DNA sequence
C	10	422.4	27.5	5000	19	AAV20609	Human kynurenine-3
	11	400.4	26.1	449	20	AAZ42057	Human endometrium

AAA38861

ID AAA38861 standard; cDNA; 1537 BP.

XX

AC XX AAA38861;

```
DT
     31-AUG-2000 (first entry)
 XX
 DE
     Human G-protein coupled receptor, HG51, coding sequence.
 XX
     Human; G-protein coupled receptor; HG51; signal transduction;
 KW
     rhodopsin receptor; obesity; type II diabetes;
 KW
 KW
     inflammatory bowel disease; constipation; diarrhoea; gene therapy; ss.
 XX
 OS
     Homo sapiens.
 XX
 FH
     Key
                   Location/Qualifiers
 FT
     CDS
                   160..1368
 FT
                   /*tag= a
 FT
                   /product= "Human HG51"
 XX
 PN
     WO200031108-A1.
 XX
 PD
     02-JUN-2000.
 XX
 PF
     18-NOV-1999;
                  99WO-US27305.
 XX
 PR
     24-NOV-1998;
                  98US-0109717.
XX
PA
     (MERI ) MERCK & CO INC.
XX
PΤ
     Liu Q, McDonald TP;
XX
     WPI; 2000-400025/34.
DR
DR
     P-PSDB; AAY98008.
XX
PT
     New DNA encoding human HG51 (a G-protein coupled receptor), useful in
     chromosomal mapping studies for identifying the chromosomal locations
PT
PT
     of the HG51 gene(s) -
XX
     Claim 1; Fig 1; 68pp; English.
PS
XX
    G protein-coupled receptors (GPCR) are important in signal transduction
CC
CC
     from the exterior to the interior of cells. Rhodopsin receptors are a
     type of GPCR which comprise a chromophore-binding pocket which is
CC
CC
     covalently linked by a protonated Schiff base to a Lys residue in
CC
     transmembrane domain 7. The present sequence is the coding sequence of
CC
     the human HG51 GPCR and is a member of the rhodopsin receptor family of
CC
    GPCRs. Due to the Lys residue and Schiff base present in HG51, it is
CC
    thought that the HG51 ligand may be a fatty-acid-like molecule. It is
CC
    also believed that agonists and antagonists of HG51 are useful for
CC
    treating various disorders such as obesity, type II diabetes,
CC
    inflammatory bowel disease, constipation or diarrhoea. In addition, the
CC
    present sequence may be used in gene therapy for the above mentioned
CC
XX
SO
    Sequence 1537 BP; 320 A; 426 C; 421 G; 370 T; 0 other;
 Query Match 100.0%; Score 1537; DB 21; Length 1537; Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1537; Conservative
                            0; Mismatches
                                            0; Indels
                                                        0; Gaps
                                                                   0;
Qу
       1 ggggccacgggggtgcgccggcgggtagcgcggggcccctcagtgcacaatggccag 60
        Db
       1 ggggccacgggggtgcgccggcgggcgggtagcgcggggcccctcagtgcacaatggccag 60
      61 agcaggeggeggagececagececacecagtgeggagegegegegageceegegeaag 120
Qу
        Db
      121 ctgagcgcctccgccaggcgcggcggcgggccatgtactcggggaaccgcagc 180
Qу
        Db
     121 ctgagcgcctccgcccggccaggcgccggcgccgggccatgtactcggggaaccgcagc 180
Qу
```

Db	181 ggcggccacggctactgggacggcggggggcggcgggggggg
Qу	241 gggacactgagccccgcgcccctcttcagccccggcacctacgagcgcctggcgctgctg 300
Db	241 gggacactgagccccgcgcccctcttcagccccggcacctacgagcgcctggcgctgctg 300
Qу	301 ctgggctccattgggctgctgggcgtcggcaacaacctgctggtgctcgtcctctactac 360
Db	301 ctgggctccattgggctgctgggcgtcggcaacaacctgctggtgctcgtcctctactac 360
Qy Db	361 aagttecageggeteegeacteceacteaceteeteggteaacateageeteagegae 420
Qy	361 aagttccagcggctccgcactcccactcactcctcctggtcaacatcagcctcagcgac 420
Db	421 ctgctggtgtccctcttcggggtcacctttaccttcgtgtcctgcctg
Qу	481 gtgtgggacaccgtgggctgcgtgtgqqacgggtttagcggcagcctcttcgggattgtt 540
Db	
Qу	541 tccattgccaccctaaccgtgctggcctatgaacgttacattcgcgtggtccatgccaga 600
Db	541 tccattgccaccctaaccgtgctggcctatgaacgttacattcgcgtggtccatgccaga 600
Qу	601 gtgatcaatttttcctgggcctggagggccattacctacatctggctctactcactggcg 660
Db	601 gtgatcaattttteetgggeetggagggeeattaeetaeatetggetetaeteaetggeg 660
Qу	661 tgggcaggagcacctctcctgggatggaacaggtacatcctggacgtacacggactaggc 720
Db	661 tgggcaggagcacctctcctgggatggaacaggtacatcctggacgtacacggactaggc 720
Qy Db	721 tgcactgtggactggaaatccaaggatgccaacgattcctcctttgtgcttttcttattt 780
Qy	721 tgcactgtggactggaaatccaaggatgccaacgattcctcctttgtgcttttcttattt 780 781 cttggctgcctggtggtgcccctgggtgtcatagcccattgctatggccatattctatat 840
Db	
Qу	841 tccattcgaatgcttcgttgtgtggaagatcttcagacaattcaagtgatcaagatttta 900
Db	
Qу	901 aaatatgaaaagaaactggccaaaatgtgctttttaatgatattcaccttcctggtctgt 960
Db	901 dadtatgaaaagaaactggccaaaatgtgctttttaatgatattcaccttcctggtctgt 960
QУ	961 tggatgccttatatcgtgatctgcttcttggtggttaatggtcatggtcacctggtcact 1020
Db	961 tggatgccttatatcgtgatctgcttcttggtggttaatggtcatggtcacctggtcact 1020
Qy Db	1021 ccaacaatatctattgtttcgtacctctttgctaaatcgaacactgtatacaatccagtg 1080
Qy	1021 ccaacaatatctattgtttcgtacctctttgctaaatcgaacactgtatacaatccagtg 1080 1081 atttatgtcttcatgatcagaaagtttcgaagatcccttttgcagcttctgtgcctccga 1140
Db	1081 atttatgtcttcatgatcagaaagtttcgaagatcccttttgcagcttctgtgcctccga 1140
Qy	1141 ctgctgaggtgccagaggcctqctaaaqacctaccagcagctggaagtgaaatgcagatc 1200
Db	
Qy	1201 agacccattgtgatgtcacagaaagatggggacaggccaaagaaaaagagaaaagtgactttcaag 1260
Db	
Qy	1261 tcttcttccatcatttttatcatcaccagtgatgaatcactgtcagttgacgacagcgac 1320

```
Db
      1261 tettettecateatttttateateaceagtgatgaateactgteagttgacgacagegac 1320
 Qу
      1321 aaaaccaatgggtccaaagttgatgtaatccaagttcgtcctttgtaggaatgaagaatg 1380
           1321 aaaaccaatgggtccaaagttgatgtaatccaagttcgtcctttgtaggaatgaagaatg 1380
 Db
 Qу
      1381 gcaacgaaagatggggccttaaattggatgccacttttggactttcatcataagaagtgt 1440
           1381 gcaacgaaagatggggccttaaattggatgccacttttggactttcatcataagaagtgt 1440
 Db
      1441 ctggaatacccgttctatgtaatatcaacagaaccttgtggtccagcaggaaatccgaat 1500
 Qy
           DЪ
      1441 ctggaatacccgttctatgtaatatcaacagaaccttgtggtccagcaggaaatccgaat 1500
     1501 tgcccatatgctcttgggcctcaggaagaggttgaac 1537
 Qу
           Db
     1501 tgcccatatgctcttgggcctcaggaagaggttgaac 1537
 RESULT
 AAA73212
     AAA73212 standard; cDNA; 2144 BP.
 XX
 AC
     AAA73212;
 XX
 DT
     05-DEC-2000 (first entry)
 XX
 DΕ
     Human 17723 receptor protein encoding cDNA SEQ ID NO:2.
 XX
 KW
     Human; 17723 receptor protein; chromosome 1q42-44; diagnosis; vaccine;
 KW
     G-protein coupled receptor; gene therapy; ss.
 XX
 os
     Homo sapiens.
 XX
 PN
     WO200043513-A1.
 XX
 PD
     27-JUL-2000.
XX
     21-JAN-2000; 2000WO-US01592.
 PF
XX
PR
     21-JAN-1999:
                  99US-0234923.
XX
     (MILL-) MILLENNIUM PHARM INC.
PA
хx
ΡI
     Glucksmann MA:
XX
DR
     WPI; 2000-476196/41.
DR
     P-PSDB; AAB12827.
XX
PT
     A G-protein-coupled receptor designated 17723 and the nucleic acids
     that encode it, useful for preventing, diagnosing and treating disorder
PΤ
PΤ
     associated with inappropriate expression of 17723 receptors -
XX
PS
    Claim 3; Page 72-73; 79pp; English.
XX
CC
    The present sequence encodes the human 17723 receptor protein (I), which
    belongs to the superfamily of G-protein-coupled receptors. (I) and the
CC
CC
    polynucleotide encoding it may be used in the prevention, treatment and
    diagnosis of diseases associated with inappropriate 17723 receptor
CC
CC
    expression. They may also be used to study the expression and function
CC
    of 17723 receptor polypeptides and their role in metabolism. The 17723
CC
    receptor polypeptides may be used as antigens in the production of
    antibodies against 17723 receptors and in assays to identify modulators
CC
CC
    (agonists and antagonists) of 17723 receptor expression and activity.
    The anti-17723 receptor antibodies and 17723 receptor antagonists may be
CC
    used to down regulate 17723 receptor expression and activity. The
CC
CC
    anti-17723 receptor antibodies may also be used as diagnostic agents for
CC
    detecting the presence of 17723 receptor polypeptides in samples
    (e.g. by enzyme linked immunosorbent assay (ELISA)). The 17723 receptor
CC
    protein has been mapped to chromosome 1q42-44.
```

1;

94.9%; Score 1459; DB 21; Length 2144; Query Match Best Local Similarity 99.9%; Pred. No. 0; Matches 1470; Conservative 0; Mismatches Indels 1; Gaps Qу 67 cggcggagccccagcccacccagtgcggagcgcgcgcgagccccgccgcaagctgagc 126 Db 1 cggcggagccccag-cccacccagtgcggagcgcgcgcggagccccgccgcaagctgagc 59 127 gcctccgcccgccaggcgccggcgccgggccatgtactcggggaaccgcagcggcggc 186 Qу Db 60 gcctccgcccgccaggcgccggcgccgggccatgtactcggggaaccgcagcggcggc 119 Qу Db 247 ctgagccccgcgcccctcttcagccccggcacctacgagcgcctggcgctgctgctgggc 306 Qу Db 180 ctgagccccgcgcccctcttcagccccggcacctacgagcgcctggcgctgctgctgctgggc 239 Qy 307 tccattgggctgctgggcgtcggcaacaacctgctggtgctcgtcctctactacaagttc 366 Db 240 tccattgggctgctgggcgtcggcaacaacctgctggtgctcgtcctctactacaagttc 299 367 cagcggctccgcactcccactcacctcctggtcaacatcagcctcagcgacctgctg 426 Qy 300 cagcggctccgcactcccactcacctcctcggtcaacatcagcctcagcgacctgctg 359 Qу Db 487 gacaccgtgggctgcgtgtgggacgggtttagcggcagcctcttcgggattgtttccatt 546 Qу Db 420 gacaccgtgggctgcgtgtgggacgggtttagcggcagcctcttcgggattgtttccatt 479 547 gccaccctaaccgtgctggcctatgaacgttacattcgcgtggtccatgccagagtgatc 606 Qу Db 480 gccaccctaaccgtgctggcctatgaacgttacattcgcgtggtccatgccagagtgatc 539 Qу 607 aatttttcctgggcctggagggccattacctacatctggctctactcactggcgtgggca 666 540 aatttttcctgggcctggagggccattacctacatctggctctactcactggcgtgggca 599 Db ggagcacctctcctgggatggaacaggtacatcctggacgtacacggactaggctgcact 726 Qy Dh 600 ggagcacctctcctgggatggaacaggtacatcctggacgtacacggactaggctgcact 659 727 gtggactggaaatccaaggatgccaacgattcctcctttgtgcttttcttatttcttggc 786 Qу 660 gtggactggaaatccaaggatgccaacgattcctcctttgtgcttttcttatttcttggc 719 Db 787 tgcctggtggtgcccctgggtgtcatagcccattgctatggccatattctatattccatt 846 Qу Db 720 tgcctggtggtgcccctgggtgtcatagcccattgctatggccatattctatattccatt 779 Qу 847 cgaatgcttcgttgtgtggaagatcttcagacaattcaagtgatcaagattttaaaatat 906 Db 780 cgaatgcttcgttgtgtggaagatcttcagacaattcaagtgatcaagattttaaaaatat 839 Qу 907 gaaaagaaactggccaaaatgtgctttttaatgatattcaccttcctggtctgttggatg 966 Db 840 gaaaagaaactggccaaaatgtgctttttaatgatattcaccttcctggtctgttggatg 899 Qy 967 ccttatatcgtgatctgcttcttggtggttaatggtcatggtcacctggtcactccaaca 1026 Dh 900 ccttatatcgtgatctgcttcttggtggttaatggtcatggtcacctggtcactccaaca 959

Qу	1027	atatctattgtttcgtacctctttgctaaatcgaacactgtatacaatccagtgatttat 1086
Db	960	atatctattgtttcgtacctctttgctaaatcgaacactgtatacaatccagtgatttat 1019
Qy	1087	gtcttcatgatcagaaagtttcgaagatcccttttgcagcttctgtgcctccgactgctg 1146
Db	1020	gtcttcatgatcagaaagtttcgaagatcccttttgcagcttctgtgcctccgactgctg 1079
Qy	1147	aggtgccagaggcctgctaaagacctaccagcagctggaagtgaaatgcagatcagaccc 1206
Db	1080	aggtgccagaggcctgctaaagacctaccagcagctggaagtgaaatgcagatcagaccc 1139
Qy	1207	attgtgatgtcacagaaagatggggacaggccaaagaaaaagtgactttcaactcttct 1266
Db	1140	attgtgatgtcacagaaagatggggacaggccaaagaaaaagtgactttcaactcttct 1199
Qy	1267	tccatcatttttatcatcaccagtgatgaatcactgtcagttgacgacagcgacaaaacc 1326
Db	1200	tccatcatttttatcatcaccagtgatgaatcactgtcagttgacgacagcgacaaaacc 1259
Qy	1327	aatgggtccaaagttgatgtaatccaagttcgtcctttgtaggaatgaagaatggcaacg 1386
Db	1260	aatgggtccaaagttgatgtaatccaagttcgtcctttgtaggaatgaagaatggcaacg 1319
Qу	1387	aaagatggggccttaaattggatgccacttttggactttcatcataagaagtgtctggaa 1446
Db	1320	aaagatggggccttaaattggatgccacttttggactttcatcataagaagtgtctggaa 1379
Qу	1447	tacccgttctatgtaatatcaacagaaccttgtggtccagcaggaaatccgaattgccca 1506
Db	1380	tacccgttctatgtaatatcaacagaaccttgtggtccagcaggaaatccgaattgccca 1439
Qу	1507	tatgetettgggeetcaggaagaggttgaac 1537
Db	1440	tatgetettgggeeteaggaagaggttgaae 1470

Res	ult No.	Score	% Query Match	Length	DB	ID	Description
С	1	422.4	27.5	5000	3	US-09-147-522-5	Sequence 5, Appli
	2	107.6	7.0	3016	1	US-07-805-123C-1	Sequence 1, Appli
	3	107.6	7.0	3016	1	US-08-033-081B-1	Sequence 1, Appli
	4	78	5.1	1105	2	US-08-466-103A-15	Sequence 15, Appl
	5	73	4.7	1410	4	US-09-255-368-1	Sequence 1, Appli
	6	69.2	4.5	1420	1	US-08-358-171-1	Sequence 1, Appli
	7	69.2	4.5	1420	3	US-09-090-947-1	Sequence 1, Appli
	8	67.2	4.4	1293	4	US-09-255-368-7	Sequence 7, Appli
	9	67	4.4	1776	1	US-08-722-001-29	Sequence 29, Appl
	10	65.4	4.3	2140	1	US-08-334-698-1	Sequence 1, Appli
	11	65.4	4.3	2140	1	US-08-228-932-1	Sequence 1, Appli
	12	65.4	4.3	2140	1	US-08-468-939-1	Sequence 1, Appli

Result No.	Score	Query Match	Length	DB	ID	Description
1	660.8	43.0	789	10	BI818538	BI818538 603033059
2	644.8	42.0	770	10	BI260681	BI260681 602968193
3	609.8	39.7	909	10	BE894106	BE894106 601438234
4	580.8	37.8	736	10	BI086726	BI086726 602850078
5	577.2	37.6	835	10	BF970560	BF970560 602274056
6	575	37.4	850	10	BI757207	BI757207 603030709
7	565.4	36.8	748	10	BG252201	BG252201 602365072
8	515.8	33.6	741	10	BG564220	BG564220 602586010
9	467.8	30.4	788	10	BF977798	BF977798 602148633
10	461.8	30.0	784	10	BI758685	BI758685 603024224

11	426.8	27.8	631	9	BB640431	BB640431 BB640431
12	426	27.7	819	10	BI088684	BI088684 602851458
13	423	27.5	424	10	BM194008	BM194008 TCAAP1E64
14	406.8	26.5	742	10	BI257225	BI257225 602976885
15	398.4	25.9	615	10	BF132059	BF132059 601821062

SEQ ID NO: 2

XX

CC

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2117	100.0	402	21	AAB12827	Human 17723 recept
2	2117	100.0	402	21	AAY98008	Human G-protein co
3	2105	99.4	402	22	AAE12070	Dendritic cell (DC
4	1063	50.2	199	22	AAB64743	Human secreted pro
5	756	35.7	147	21	AAY32195	Human receptor mol
6	664	31.4	163	22	AAE12071	Dendritic cell (DC
7	664	31.4	879	22	AAU31008	Novel human secret
8	572	27.0	123	20	AAY60172	Human endometrium
9	564	26.6	122	21	AAB38327	Human secreted pro
10	459.5	21.7	349	10	AAP90554	Bovine rhodopsin.
11	455	21.5	348	17	AAR93116	Rhodopsin. Homo s
12	451	21.3	348	21	AAY98009	Human rhodopsin re
13	449	21.2	348	14	AAR38483	Rhodopsin protein.
14	424	20.0	354	21	AAY57086	Rhodopsin amino ac
15	420.5	19.9	309	15	AAR48735	G-protein coupled
16	420.5	19.9	309	17	AAW02707	G-protein coupled

```
RESULT 1
AAB12827
     AAB12827 standard; Protein; 402 AA.
XX
AC
     AAB12827;
XX
     05-DEC-2000 (first entry)
DT
XX
     Human 17723 receptor protein SEQ ID NO:1.
DΕ
XX
KW
     Human; 17723 receptor protein; chromosome 1q42-44; diagnosis; vaccine;
KW
     G-protein coupled receptor; gene therapy.
XX
os
     Homo sapiens.
XX
PN
     WO200043513-A1.
XX
PD
     27-JUL-2000.
XX
PF
     21-JAN-2000; 2000WO-US01592.
XX
PR
     21-JAN-1999; 99US-0234923.
XX
PA
     (MILL-) MILLENNIUM PHARM INC.
XX
ΡI
     Glucksmann MA;
xx
DR
     WPI; 2000-476196/41.
DR
     N-PSDB; AAA73212.
XX
PT
     A G-protein-coupled receptor designated 17723 and the nucleic acids
PT
     that encode it, useful for preventing, diagnosing and treating disorder
     associated with inappropriate expression of 17723 receptors -
PT
XX
    Claim 1; Page 70-72; 79pp; English.
PS
```

The present sequence is the human 17723 receptor protein (I), which

```
belongs to the superfamily of G-protein-coupled receptors. (I) and the
     polynucleotide encoding it may be used in the prevention, treatment and
 CC
     diagnosis of diseases associated with inappropriate 17723 receptor
     expression. They may also be used to study the expression and function
 CC
     of 17723 receptor polypeptides and their role in metabolism. The 17723
 CC
     receptor polypeptides may be used as antigens in the production of
     antibodies against 17723 receptors and in assays to identify modulators
 CC
     (agonists and antagonists) of 17723 receptor expression and activity.
 CC
     The anti-17723 receptor antibodies and 17723 receptor antagonists may be
     used to down regulate 17723 receptor expression and activity. The
 CC
     anti-17723 receptor antibodies may also be used as diagnostic agents for
     detecting the presence of 17723 receptor polypeptides in samples
 CC
     (e.g. by enzyme linked immunosorbent assay (ELISA)). The 17723 receptor
CC
     protein has been mapped to chromosome 1q42-44.
 XX
 SO
     Sequence
              402 AA;
  Query Match
                       100.0%; Score 2117; DB 21; Length 402;
  Best Local Similarity 100.0%; Pred. No. 6.3e-222;
  Matches 402; Conservative
                             0; Mismatches
                                            0; Indels
                                                         0; Gaps
                                                                    0;
       1 MYSGNRSGGHGYWDGGGAAGAEGPAPAGTLSPAPLFSPGTYERLALLLGSIGLLGVGNNL 60
         {\tt 1~mysgnrsgghgywdgggaagaegpapagtlspaplfspgtyerlalllgsigllgvgnnl~60}\\
Db
      61 LVLVLYYKFQRLRTPTHLLLVNISLSDLLVSLFGVTFTFVSCLRNGWVWDTVGCVWDGFS 120
0v
         Dh
      61 lvlvlyykfqrlrtpthlllvnislsdllvslfgvtftfvsclrngwvwdtvgcvwdgfs 120
      121 GSLFGIVSIATLTVLAYERYIRVVHARVINFSWAWRAITYIWLYSLAWAGAPLLGWNRYI 180
Qу
         121 gslfgivsiatltvlayeryirvvharvinfswawraityiwlyslawagapllgwnryi 180
Db
Qу
     181 LDVHGLGCTVDWKSKDANDSSFVLFLFLGCLVVPLGVIAHCYGHILYSIRMLRCVEDLQT 240
         Db
     181 ldvhglgctvdwkskdandssfvlflflgclvvplgviahcyghilysirmlrcvedlqt 240
     241 IQVIKILKYEKKLAKMCFLMIFTFLVCWMPYIVICFLVVNGHGHLVTPTISIVSYLFAKS 300
Qу
         241 iqvikilkyekklakmcflmiftflvcwmpyivicflvvnghghlvtptisivsylfaks 300
Db
     301 NTVYNPVIYVFMIRKFRRSLLQLLCLRLLRCQRPAKDLPAAGSEMQIRPIVMSQKDGDRP 360
Qу
         301 ntvynpviyvfmirkfrrsllqllclrllrcqrpakdlpaagsemqirpivmsqkdgdrp 360
Dh
     361 KKKVTFNSSSIIFIITSDESLSVDDSDKTNGSKVDVIQVRPL 402
Qу
         361 kkkvtfnsssiifiitsdeslsvddsdktngskvdviqvrpl 402
Dh
RESULT
AAY98008
ID
    AAY98008 standard; Protein; 402 AA.
XX
AC
    AAY98008:
XX
DT
    31-AUG-2000 (first entry)
XX
DE
    Human G-protein coupled receptor, HG51.
XX
KW
    Human; G-protein coupled receptor; HG51; signal transduction;
KW
    rhodopsin receptor; obesity; type II diabetes;
    inflammatory bowel disease; constipation; diarrhoea; gene therapy.
KW
XX
os
    Homo sapiens.
хx
PN
    WO200031108-A1.
ХX
PD
    02-JUN-2000.
XX
```

```
PF
     18-NOV-1999;
                  99WO-US27305.
 XX
 PR
     24-NOV-1998;
                  98US-0109717.
 XX
 PA
     (MERI ) MERCK & CO INC.
 XX
     Liu Q, McDonald TP;
 ХХ
 DR
     WPI; 2000-400025/34.
 DR
     N-PSDB; AAA38861.
 PT
     New DNA encoding human HG51 (a G-protein coupled receptor), useful in
     chromosomal mapping studies for identifying the chromosomal locations
 PT
     of the HG51 gene(s) -
 PT
 XX
 PS
     Claim 23; Fig 2; 68pp; English.
 XX
 CC
     G protein-coupled receptors (GPCR) are important in signal transduction
     from the exterior to the interior of cells. Rhodopsin receptors are a
CC
     type of GPCR which comprise a chromophore-binding pocket which is
CC
     covalently linked by a protonated Schiff base to a Lys residue in
CC
     transmembrane domain 7. The present sequence is the human HG51 GPCR and
CC
     is a member of the rhodopsin receptor family of GPCRs. Due to the Lys
CC
CC
     residue and Schiff base present in HG51, it is thought that the HG51
CC
     ligand may be a fatty-acid-like molecule. It is also believed that
     agonists and antagonists of HG51 are useful for treating various
CC
     disorders such as obesity, type II diabetes, inflammatory bowel disease,
     constipation or diarrhoea. In addition, the coding sequence for the
CC
CC
     present sequence may be used in gene therapy for the above mentioned
CC
     disorders.
XX
SQ
     Sequence
              402 AA;
  Query Match 100.0%; Score 2117; DB 21; Length 402; Best Local Similarity 100.0%; Pred. No. 6.3e-222;
  Matches 402; Conservative
                             0; Mismatches
                                             0; Indels
                                                         0; Gaps
                                                                    0;
Qу
       1 MYSGNRSGGHGYWDGGGAAGAEGPAPAGTLSPAPLFSPGTYERLALLLGSIGLLGVGNNL 60
         Db
       {\tt 1~mysgnrsgghgywdgggaagaegpapagtlspaplfspgtyerlalllgsigllgvgnnl~60}\\
Οv
      61 LVLVLYYKFQRLRTPTHLLLVNISLSDLLVSLFGVTFTFVSCLRNGWVWDTVGCVWDGFS 120
         61 lvlvlyykfqrlrtpthlllvnislsdllvslfgvtftfvsclrngwvwdtvgcvwdgfs 120
Qу
     121 GSLFGIVSIATLTVLAYERYIRVVHARVINFSWAWRAITYIWLYSLAWAGAPLLGWNRYI 180
         Db
     121 gslfgivsiatltvlayeryirvvharvinfswawraityiwlyslawagapllgwnryi 180
     181 LDVHGLGCTVDWKSKDANDSSFVLFLFLGCLVVPLGVIAHCYGHILYSIRMLRCVEDLQT 240
Qy
         Db
     181 ldvhglgctvdwkskdandssfvlflflgclvvplgviahcyghilysirmlrcvedlqt 240
Оv
     241 IQVIKILKYEKKLAKMCFLMIFTFLVCWMPYIVICFLVVNGHGHLVTPTISIVSYLFAKS 300
         241 iqvikilkyekklakmcflmiftflvcwmpyivicflvvnghghlvtptisivsylfaks 300
Db
     301 NTVYNPVIYVFMIRKFRRSLLQLLCLRLLRCQRPAKDLPAAGSEMQIRPIVMSQKDGDRP 360
Qу
         301 ntvynpviyvfmirkfrrsllqllclrllrcqrpakdlpaagsemqirpivmsqkdgdrp 360
Db
Qy
     361 KKKVTFNSSSIIFIITSDESLSVDDSDKTNGSKVDVIQVRPL 402
         Db
     361 kkkvtfnsssiifiitsdeslsvddsdktngskvdviqvrpl 402
RESULT
```

AAE12070

ID AAE12070 standard; Protein; 402 AA.

```
AC
      AAE12070:
 XX
 DT
      18-DEC-2001 (first entry)
 XX
 DE
      Dendritic cell (DC) DCEPR protein.
 XX
      Dendritic cell; DC; DCEPR protein; gene therapy; dermatological; vaccine;
 KW
      atopic dermatitis; autoimmune disease; inflammatory skin disease; cancer;
      immunosuppressive; AIDS; Acquired immune deficiency syndrome; cytostatic;
 KW
 KW
      chromosomal identification; pharmaceutical; hypersensitivity; virucide;
      transplant rejection; chronic inflammatory disease; anti-HIV.
 KW
 XX
 os
      Unidentified.
 XX
 PN
      WO200172773-A2.
 XX
 PD
      04-OCT-2001.
 XX
 PF
      28-MAR-2001; 2001WO-EP03542.
 XX
 PR
     29-MAR-2000; 2000US-192934P.
     18-MAY-2000; 2000US-205020P.
 PR
 PR
     18-MAY-2000; 2000US-205026P.
     19-MAY-2000; 2000US-205767P.
 PR
 PR
     19-MAY-2000; 2000US-205769P.
 XX
 PA
      (NOVS ) NOVARTIS AG.
      (NOVS ) NOVARTIS-ERFINDUNGEN VERW GES MBH.
 PΑ
XX
 PΙ
     Werner G, Phares W, Jaritz M, Lapp H, Kalthoff FS;
XX
DR
     WPI; 2001-616466/71.
DR
     N-PSDB; AAD19720.
XX
PT
     New polypeptides for screening therapeutic agonists and antagonists
     comprise dendritic cell polypeptides -
PΤ
XX
     Claim 2; Page 42; 52pp; English.
PS
XX
CC
     The invention relates to dendritic cell (DC) proteins and their
CC
     corresponding DNA molecules. A pharmaceutical composition comprising
CC
     agonist and antagonist of DC proteins are useful for treating abnormal
CC
     conditions related to both an excess of and insufficient level of
CC
     expression of DC gene, or related to both an excess of and insufficient
CC
     activity of DC protein. Soluble form of DC proteins are used as an active
     ingredient in combination with pharmaceutical acceptable carriers.
CC
     DC genes and proteins are useful for treating chronic inflammatory
CC
CC
     diseases, autoimmune diseases, transplant rejection crisis, including
CC
     inflammatory skin diseases such as contact hypersensitivity, atopic
     dermatitis or virally-induced immune suppression such as AIDS and cancer.
CC
CC
     DC protein is useful for inducing immunological response in a mammal, and
     as immunogen to produce antibodies immunospecific for the polypeptide.
CC
CC
     DC gene is useful in gene therapy. DC gene is also useful as a diagnostic
CC
     reagent, and for chromosomal identification. The present sequence is
     dendritic cell (DC) DCEPR protein which is found to belong to the family
CC
CC
    of G-protein coupled receptor protein.
XX
SQ
    Sequence
               402 AA:
  Query Match
                         99.4%; Score 2105; DB 22;
                                                     Length 402;
  Best Local Similarity 99.5%; Pred. No. 1.3e-220;
 Matches 400; Conservative
                              1; Mismatches
                                                 1:
                                                     Indels
                                                                          0;
       1 MYSGNRSGGHGYWDGGGAAGAEGPAPAGTLSPAPLFSPGTYERLALLLGSIGLLGVGNNL 60
Qу
         Db
       {\tt 1~mysgnrsgghgywdgggaagakgpapagtlspaplfspgtyerlalllgsigllgvgnnl~60}\\
0v
      61 LVLVLYYKFQRLRTPTHLLLVNISLSDLLVSLFGVTFTFVSCLRNGWVWDTVGCVWDGFS 120
         61 lvlvlyykfqrlrtpthlllvnislsdllvslfgvtftfvsclrngwvwdtvgcvwdgfs 120
```

```
121 GSLFGIVSIATLTVLAYERYIRVVHARVINFSWAWRAITYIWLYSLAWAGAPLLGWNRYI 180
 Qу
           121 gslfgivsiatltvlayeryirvvharvinfswawraityiwlyslawagapllgwnryi 180
 Db
       181 LDVHGLGCTVDWKSKDANDSSFVLFLFLGCLVVPLGVIAHCYGHILYSIRMLRCVEDLQT 240
 Qу
           181 ldvhglgctvdwkskdandssfvlflflgclvvplgviahcyghilysirmlrcvedlqt 240
 Db
       241 IQVIKILKYEKKLAKMCFLMIFTFLVCWMPYIVICFLVVNGHGHLVTPTISIVSYLFAKS 300
 Qy
           111441114411114411
 Db
       241 iqvikilkyekklakmcflmiftflvcwmpyivicflvvnghghlvtptisivsylfaks 300
       301 NTVYNPVIYVFMIRKFRRSLLQLLCLRLLRCQRPAKDLPAAGSEMQIRPIVMSQKDGDRP 360
 Qγ
           1144411441141141414141414141414
       301 ntvynpviyvfmirkfrrsllqllclrllrcqrpakdlpaagsemqirpivmsqkdgdrp 360
 Db
       361 KKKVTFNSSSIIFIITSDESLSVDDSDKTNGSKVDVIQVRPL 402
 Оy
          Db
       361 kkkvtfnsssiifigtsdeslsvddsdktngskvdviqvrpl 402
 RESULT
 AAB64743
     AAB64743 standard; Protein; 199 AA.
 AC
     AAB64743;
 XX
 DΤ
     23-MAR-2001 (first entry)
 XX
 DE
     Human secreted protein sequence encoded by gene 15 SEQ ID NO:137.
 XX
 KW
     Human; secreted protein; diagnosis; cytostatic; antirheumatic;
 KW
     antiarthritic; dermalogical; cardiant; antiinflammatory; anti-ulcer;
     gastrointestinal; solid tumour; rheumatoid arthritis; psoriasis;
 KW
     diabetic retinopathy; myocardial angiogenesis; Crohn's disease;
 KW
 KW
     ulcer.
 XX
 OS
     Homo sapiens.
 XX
 PN
     WO200077237-A1.
 ХX
 PD
     21-DEC-2000.
XX
 ΡF
     01-JUN-2000; 2000WO-US14928.
XX
PR
     11-JUN-1999;
                  99US-0138633.
XX
     (HUMA-) HUMAN GENOME SCI INC.
PA
PΑ
     (ROSE/) ROSEN C A.
XX
ΡI
     Rosen CA, Ruben SM, Komatsoulis GA;
XX
DR
     WPI; 2001-071280/08.
XX
PT
    Nucleic acids encoding 49 human secreted polypeptides, useful for
РΤ
     preventing, diagnosing and/or treating diseases such as tumors,
PT
     rheumatoid arthritis, psoriasis and diabetic retinopathy -
XX
PS
    Disclosure; Page 503; 520pp; English.
XX
CC
    The polynucleotide sequences given in AAF33037 to AAF33085 encode the
CC
    human secreted proteins given in AAB64666 to AAB64714. AAB64715 to
    AAB64771 represent human secreted polypeptide sequences and proteins
CC
CC
    homologous to them, which are given in the exemplification of the present
CC
    invention. Human secreted proteins have activities based on the tissues
CC
    and cells the genes are expressed in. Examples of activities include:
CC
    cytostatic; antirheumatic; antiarthritic; dermalogical; cardiant;
CÇ
    antiinflammatory; gastrointestinal; and anti-ulcer. The polynucleotides
    and polypeptides can be used in the prevention, treatment and diagnosis
CC
    of diseases associated with inappropriate polypeptide expression.
```

Disorders that may be treated or prevented include solid tumours, CC rheumatoid arthritis, psoriasis, diabetic retinopathy, myocardial CC angiogenesis, Crohn's disease and ulcers. The polynucleotides and their CC CC complementary sequences may also be used as DNA probes in diagnostic assays (e.g. polymerase chain reactions (PCR)) to detect and quantitate CC the presence of similar nucleic acid sequences in samples, and therefore CC which patients may be in need of restorative therapy. The polypeptides CC CC may also be used as antigens in the production of antibodies against the CC polypeptide and in assays to identify modulators (agonists and antagonists) of polypeptide expression and activity. The anti-polypeptide CC antibodies and antagonists may also be used to down regulate expression CC CC and activity. AAF33028 to AAF33036 and AAB64665 represent sequences used CC in the exemplification of the present invention. XX SQ Sequence 199 AA; Query Match 50.2%; Score 1063; DB 22; Length 199; Best Local Similarity 100.0%; Pred. No. 2e-107; Matches 199; Conservative 0; Mismatches 0; Indels 0; Gaps 0; 118 GFSGSLFGIVSIATLTVLAYERYIRVVHARVINFSWAWRAITYIWLYSLAWAGAPLLGWN 177

Db 1 gfsgslfgivsiatltvlayeryirvvharvinfswawraityiwlyslawagapllgwn 60 178 RYILDVHGLGCTVDWKSKDANDSSFVLFLFLGCLVVPLGVIAHCYGHILYSIRMLRCVED 237 Qу 61 ryildvhglgctvdwkskdandssfvlflflgclvvplgviahcyghilysirmlrcved 120 Db 238 LQTIQVIKILKYEKKLAKMCFLMIFTFLVCWMPYIVICFLVVNGHGHLVTPTISIVSYLF 297 Qy 121 lqtiqvikilkyekklakmcflmiftflvcwmpyivicflvvnghghlvtptisivsylf 180 Db Qу 298 AKSNTVYNPVIYVFMIRKF 316 Db 181 aksntvynpviyvfmirkf 199

		*				
Result		Query				
No.	Score	Match	Length	DB	ID	Description
1	451	21.3	348	2	US-08-390-000A-8	Sequence 8, Appli
2	444	21.0	348	4	US-08-430-286A-11	Sequence 11, Appl
3	420.5	19.9	309	1	US-08-118-270-56	Sequence 56, Appl
4	420.5	19.9	309	5	PCT-US93-08528-56	Sequence 56, Appl
5	354.5	16.7	297	1	US-08-118-270-58	Sequence 58, Appl
6	354.5	16.7	297	5	PCT-US93-08528-58	Sequence 58, Appl
7	341.5	16.1	305	1	US-08-118-270-59	Sequence 59, Appl
8	341.5	16.1	305	5	PCT-US93-08528-59	Sequence 59, Appl
9	338.5	16.0	297	1	US-08-118-270-57	Sequence 57, Appl
10	338.5	16.0	297	5	PCT-US93-08528-57	Sequence 57, Appl
11	309	14.6	391	1	US-07-816-283-2	Sequence 2, Appli
12	309	14.6	391	1	US-08-417-103-2	Sequence 2, Appli
13	309	14.6	391	1	US-08-417-103-14	Sequence 14, Appl
14	304	14.4	391	1	US-07-816-283-4	Sequence 4, Appli
15	304	14.4	391	1	US-08-417-103-4	Sequence 4, Appli

Result No.	Score	Query Match	Length	DB	ID	Description
1	477.5	22.6	349	1	JC5490	opsin, pineal glan
2	475	22.4	351	1	A55962	opsin, pineal glan
3	464	21.9	352	2	I50081	rhodopsin - green
4	458	21.6	348	1	OOBO	rhodopsin - bovine
5	456.5	21.6	348	1	JC4267	opsin - rabbit
6	455	21.5	348	1	S23398	rhodopsin - Chines
7	452.5	21.4	351	2	S29152	rhodopsin - chicke

```
348 1 OOHU
       8
             451
                  21.3
                                                              rhodopsin - human
             451 21.3
                           354 1 S27231
348 1 A23665
       9
                                                              rhodopsin - northe
             450 21.3
                                                              opsin - mouse
             448 21.2 354 1 I51200
      11
                                                              rhodopsin - Africa
 Result
                  Ouerv
    No. Score Match Length DB ID
                                                             Description
   .....
     1 2117 100.0 402 1 OPN3_HUMAN
                                                 Q9h1y3 homo sapien
Q9wuk7 mus musculu
P51476 columba liv
P51475 gallus gall
P52202 alligator m
      2
          1862 88.0 400 1 OPN3_MOUSE
         477.5 22.6 349 1 OPSP_COLLI
475 22.4 351 1 OPSP_CHICK
468 22.1 352 1 OPSD_ALLMI
      3
      4
                                                           P52202 alligator m
      5
         466.5 22.0 444 1 OPSP_PETMA
      6
                                                             042490 petromyzon
         464 21.9 352 1 OPSD_ANOCA
458 21.6 348 1 OPSD_BOVIN
456.5 21.6 348 1 OPSD_RABIT
      7
                                                            P41591 anolis caro
      8
                                                           P02699 bos taurus
      9
                                                           P49912 oryctolagus
P28681 cricetulus
     10
         455 21.5 348 1 OPSD_CRIGR
     11
           455 21.5 348 1 OPSD_MACFA
455 21.5 354 1 OPSD_RANCA
                                                            Q28886 macaca fasc
     12
                                                            P51470 rana catesb
 RESULT
         1
 OPN3 HUMAN
     OPN3 HUMAN
                     STANDARD;
                                     PRT; 402 AA.
      Q9H1Y3; Q9Y344;
      16-OCT-2001 (Rel. 40, Created)
 DT
      16-OCT-2001 (Rel. 40, Last sequence update)
     01-MAR-2002 (Rel. 41, Last annotation update)
     Opsin 3 (Encephalopsin) (Panopsin).
 DE
 GN
      OPN3 OR ECPN.
 OS
      Homo sapiens (Human).
 OC
      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
      Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC
 OX
      NCBI_TaxID=9606;
 RN
      [1]
RP
      SEQUENCE FROM N.A.
      MEDLINE=99252448; PubMed=10234000;
 RX
RA
      Blackshaw S., Snyder S.H.;
RT
      "Encephalopsin: a novel mammalian extraretinal opsin discretely
RT
     localized in the brain.";
RL
     J. Neurosci. 19:3681-3690(1999).
RN
     [2]
RP
     SEQUENCE FROM N.A.
RX
     MEDLINE=21295039; PubMed=11401433;
     Halford S., Freedman M.S., Bellingham J., Inglis S.L.,
     Poopalasundaram S., Soni B.G., Foster R.G., Hunt D.M.;
RA
RT
     "Characterization of a novel human opsin gene with wide tissue
RТ
     expression and identification of embedded and flanking genes on
     chromosome 1q43.";
RT
RL
     Genomics 72:203-208(2001).
RN
     [3]
RP
     SEQUENCE FROM N.A.
     Parker A.:
RI,
     Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.
     -!- FUNCTION: May play a role in encephalic photoreception.
CC
CC
     -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC
     -!- TISSUE SPECIFICITY: Strongly expressed in brain. Highly expressed
CC
         in the preoptic area and paraventricular nucleus of the
CC
         hypothalamus. Shows highly patterned expression in other regions
CC
         of the brain, being enriched in selected regions of the cerebral
CC
         cortex, cerebellar Purkinje cells, a subset of striatal neurons,
CC
         selected thalamic nuclei, and a subset of interneurons in the
CC
         ventral horn of the spinal cord.
     -!- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.
CC
CC
         OPSIN SUBFAMILY.
CC
CC
```

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

```
CC
     the European Bioinformatics Institute. There are no restrictions on its
     use by non-profit institutions as long as its content is in no way
 CC
     modified and this statement is not removed. Usage by and for commercial
     entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC
     or send an email to license@isb-sib.ch).
 DR
     EMBL; AF140242; AAD32671.1; -.
     EMBL; AF303588; AAK37447.1; -.
 DR
 DR
     EMBL; AL133390; CAC19785.1; -.
     InterPro; IPR000276; GPCR_Rhodpsn.
 DR
 DR
     Pfam; PF00001; 7tm 1; 1.
 DR
     PRINTS; PR00237; GPCRRHODOPSN.
     PROSITE; PS00237; G_PROTEIN_RECEP_F1_1; 1.
 DR
     PROSITE; PS50262; G_PROTEIN_RECEP_F1_2; 1.
 DR
 DR
     PROSITE; PS00238; OPSIN; 1.
 KW
     Photoreceptor; Retinal protein; Transmembrane; Lipoprotein; Palmitate;
 KW
     G-protein coupled receptor.
 FT
     DOMATN
                1
                      40
                               EXTRACELLULAR (POTENTIAL).
 FT
     TRANSMEM
                41
                      65
                              1 (POTENTIAL)
 FT
     DOMAIN
                66
                     77
                              CYTOPLASMIC (POTENTIAL).
 FT
     TRANSMEM
                78
                     102
                              2 (POTENTIAL).
 FΤ
     DOMAIN
               103
                     117
                              EXTRACELLULAR (POTENTIAL).
     TRANSMEM
               118
                   137
                              3 (POTENTIAL)
 FT
     DOMAIN
               138
                     153
                              CYTOPLASMIC (POTENTIAL).
 FT
     TRANSMEM
               154
                     177
                              4 (POTENTIAL).
FT
     DOMATN
               178
                     201
                              EXTRACELLULAR (POTENTIAL).
     TRANSMEM
               202
                     229
                              5 (POTENTIAL).
FΤ
     DOMAIN
               230
                     255
                              CYTOPLASMIC (POTENTIAL).
FT
     TRANSMEM
               256
                     279
                              6 (POTENTIAL).
FT
     DOMAIN
               280
                     287
                              EXTRACELLULAR (POTENTIAL).
FT
     TRANSMEM
               288
                     312
                              7 (POTENTIAL).
FT
     DOMAIN
               313
                              CYTOPLASMIC (POTENTIAL).
                     402
FT
     DISULFID
               114
                     188
                              BY SIMILARITY.
FΤ
     BINDING
               299
                     299
                              RETINAL CHROMOPHORE.
FT
     LIPID
               325
                     325
                              PALMITATE (BY SIMILARITY).
FT
                              N-LINKED (GLCNAC. . .) (POTENTIAL).
     CARBOHYD
               5
                      - 5
FT
     CARBOHYD
                   198
               198
                              N-LINKED (GLCNAC. . .) (POTENTIAL).
FT
     CONFLICT
               390
                              NGSKVDV -> IGVQSLML (IN REF. 1).
                     396
     SEQUENCE 402 AA; 44873 MW; 370F64C19F834A71 CRC64;
  Query Match 100.0%; Score 2117; DB 1; Length 402; Best Local Similarity 100.0%; Pred. No. 1e-136;
  Matches 402; Conservative
                             0; Mismatches
                                            0; Indels
       1 MYSGNRSGGHGYWDGGGAAGAEGPAPAGTLSPAPLFSPGTYERLALLLGSIGLLGVGNNL 60
Qу
         1 MYSGNRSGGHGYWDGGGAAGAEGPAPAGTLSPAPLFSPGTYERLALLLGSIGLLGVGNNL 60
Db
      61 LVLVLYYKFQRLRTPTHLLLVNISLSDLLVSLFGVTFTFVSCLRNGWVWDTVGCVWDGFS 120
Qу
         61 LVLVLYYKFQRLRTPTHLLLVNISLSDLLVSLFGVTFTFVSCLRNGWVWDTVGCVWDGFS 120
Db
     121 GSLFGIVSIATLTVLAYERYIRVVHARVINFSWAWRAITYIWLYSLAWAGAPLLGWNRYI 180
Qу
         121 GSLFGIVSIATLTVLAYERYIRVVHARVINFSWAWRAITYIWLYSLAWAGAPLLGWNRYI 180
Db
     181 LDVHGLGCTVDWKSKDANDSSFVLFLFLGCLVVPLGVIAHCYGHILYSIRMLRCVEDLQT 240
0y
         181 LDVHGLGCTVDWKSKDANDSSFVLFLFLGCLVVPLGVIAHCYGHILYSIRMLRCVEDLQT 240
Db
     241 IQVIKILKYEKKLAKMCFLMIFTFLVCWMPYIVICFLVVNGHGHLVTPTISIVSYLFAKS 300
Qу
         241 IQVIKILKYEKKLAKMCFLMIFTFLVCWMPYIVICFLVVNGHGHLVTPTISIVSYLFAKS 300
Db
Qу
     301 NTVYNPVIYVFMIRKFRRSLLQLLCLRLLRCQRPAKDLPAAGSEMQIRPIVMSQKDGDRP 360
         Db
     301 NTVYNPVIYVFMIRKFRRSLLQLLCLRLLRCQRPAKDLPAAGSEMQIRPIVMSQKDGDRP 360
     361 KKKVTFNSSSIIFIITSDESLSVDDSDKTNGSKVDVIQVRPL 402
Qу
```

Result No.	Score	% Query Match	Length	DB	ID	Description
1	500.5	23.6	352	13	Q9W6K3	Q9w6k3 anolis caro
2	491.5	23.2	534	13	057422	057422 xenopus lae
3	484.5	22.9	346	13	Q9PUA9	Q9pua9 bufo japoni
4	480	22.7	357	13	Q9IBH2	Q9ibh2 phelsuma ma
5	473.5	22.4	377	13	Q9IB88	Q9ib88 brachydanio
6	473	22.3	543	13	Q90YK6	Q90yk6 gallus gall
7	458	21.6	348	6	Q95KU1	Q95kul felis silve
8	457.5	21.6	351	13	Q9IA36	Q9ia36 poephila qu
9	455.5	21.5	351	13	Q9W6S0	Q9w6s0 columba liv
10	455	21.5	363	13	Q98TH3	Q98th3 cynops pyrr
11	453.5	21.4	322	13	057448	O57448 anas platyr